PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : B60S 1/08 — A1

(11) International Publication Number: **WO 99/55564**

(43) International Publication Date: 4 November 1999 (04.11.99)

(21) International Application Number: PCT/US99/08921

(22) International Filing Date: 23 April 1999 (23.04.99)

(30) Priority Data:
60/083,025 24 April 1998 (24.04.98) US

(71) Applicant *(for all designated States except US)*: UNIVERSITY OF FLORIDA [US/US]; 223 Grinter Hall, Gainesville, FL 32611 (US).

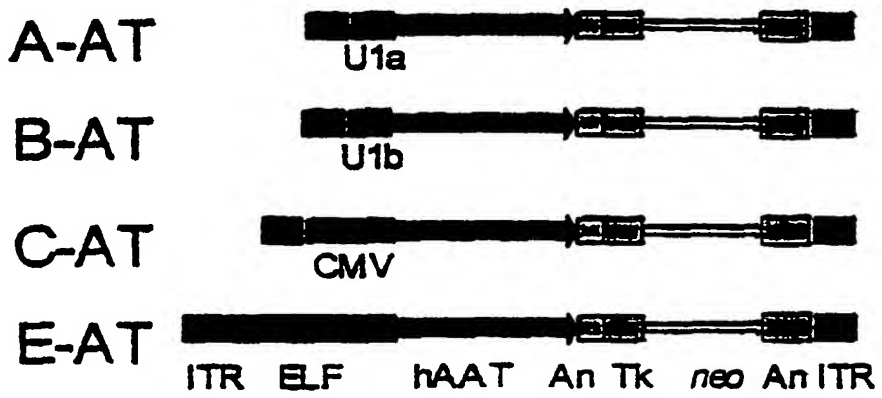
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: FLOTTE, Terence, R. [US/US]; 2058 S.W. 72nd Street, Gainesville, FL 32607 (US). SONG, Sihong [CN/US]; 3418 N.W. 67th Avenue, Gainesville, FL 32653 (US). BYRNES, Barry, J. [US/US]; 7902 S.W. 45th Lane, Gainesville, FL 32608 (US). MORGAN, Michael [US/US]; 2300 S.W. 43rd Street, K-5, Gainesville, FL 32607 (US).

(74) Agent: HIBLER, David, W.; Williams, Morgan & Amerson, P.C., Suite 250, 7676 Hillmont, Houston, TX 77040 (US).

(81) Designated States: AE, AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: MATERIALS AND METHODS FOR GENE THERAPY



(57) Abstract

The subject invention concerns materials and methods for gene therapy. One aspect of the invention pertains to vectors which can be used to effect genetic therapy in animals or humans having genetic disorders where expression of high levels of a protein of interest are required to treat or correct the disorder. The subject invention also pertains to methods for treating animals or humans in need of gene therapy to treat or correct a genetic disorder. The materials and methods of the invention can be used to provide therapeutically effective levels of a protein that is non-functional, or that is absent or deficient in the animal or human to be treated. In one embodiment, the materials and methods can be used to treat alpha-1-antitrypsin deficiency.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

DESCRIPTION

MATERIALS AND METHODS FOR GENE THERAPY

The subject invention was made with government support under a research project supported by National Institute of Health NHLBI Grant No. HL 59412. The government has certain rights in this invention.

Cross-Reference to a Related Application

This application claims priority from provisional application U.S. Serial No. 60/083,025, filed April 24, 1998.

Background of the Invention

Alpha-1-antitrypsin (AAT) deficiency is the second most common monogenic lung disease in man, accounting for approximately 3% of all early deaths due to obstructive pulmonary disease. AAT protein is normally produced in the liver, secreted into the serum and circulated to the lung where it protects the fine supporting network of elastin fibers from degradation by neutrophil elastase. Current therapy for AAT deficiency includes avoidance of cigarette smoke exposure and weekly intravenous infusions of recombinant human AAT (hAAT) protein. Attempts to devise gene therapy strategies to replace AAT either in the lung itself or within any of a number of other tissues which are capable of AAT secretion have been limited by the short duration of expression from some vectors and by the relatively high circulating levels of AAT which is required for therapeutic effect. Methods of gene therapy have been described in U.S. Patent No. 5,399,346.

It has recently been demonstrated that adeno-associated virus (AAV) vectors are capable of stable *in vivo* expression and may be less immunogenic than other viral vectors (Flotte *et al.*, 1996; Xiao *et al.*, 1996; Kessler *et al.*, 1996; Jooss *et al.*, 1998). AAV is a non-pathogenic human parvovirus whose life cycle naturally includes a mechanism for long-term latency. In the case of wild-type AAV (wtAAV), this persistence is due to site-specific integration into a site on human chromosome 19 (the AAVSI site) in the majority of cells (Kotin *et al.*, 1990), whereas with recombinant

AAV (rAAV) vectors, persistence appears to be due to a combination of episomal persistence and integration into non-chromosome 19 locations (Afione *et al.*, 1996; Kearns *et al.*, 1996). Recombinant AAV latency also differs from that of wtAAV in that wtAAV is rapidly converted to double-stranded DNA in the absence of helper virus (*e.g.*, adenovirus) infection, while with rAAV leading strand synthesis is delayed in the absence of helper virus (Fisher *et al.*, 1996; Ferrari *et al.*, 1996). U.S. Patent No. 5,658,785 describes adeno-associated virus vectors and methods for gene transfer to cells.

Kessler *et al.* (1996) demonstrated that murine skeletal myofibers transduced by an rAAV vector were capable of sustained secretion of biologically active human erythropoietin (hEpo), apparently without eliciting a significant immune response against the secreted hEpo. See also U.S. Patent No. 5,858,351 issued to Podsakoff *et al.* Likewise, Murphy *et al.* (1997) have observed the expression and secretion of sustained levels of leptin in *ob/ob* mice after AAV muscle transduction. Brantly *et al.* (U.S. Patent No. 5,439,824) disclose methods for increasing expression of AAT using vectors comprising intron II of the human AAT gene. However, the level of leptin expression observed was only in the range of 2 to 5 ng/ml. Therapy for AAT deficiency requires serum levels of at least about 800 μg/ml. Thus, there remains a need in the art for a means of providing therapeutically beneficial levels of a protein to a person in need of such treatment.

Brief Summary of the Invention

The subject invention concerns materials and methods for gene therapy. One aspect of the invention pertains to vectors which can be used to provide genetic therapy in animals or humans having a genetic disorder where relatively high levels of expression of a protein is required to treat the disorder. The vectors of the invention are based on adeno-associated virus (AAV). The vectors are designed to provide high levels of expression of heterologous DNA contained in the vector. In one embodiment, the vectors comprise AAV inverted terminal repeat sequences and constitutive or regulatable promoters for driving high levels of gene expression. The subject invention also pertains to methods for treating animals or humans in need of gene therapy, *e.g.*, to correct a genetic deficiency disorder.

Brief Description of the Drawings

Figure 1 shows rAAV-AAT vector cassettes used according to the subject invention. The A-AT and B-AT constructs contain the promoters from the small nuclear RNA genes, U1a and U1b, respectively. The C-AT construct contains the CMV promoter, whereas the E-AT vector uses the human elongation factor 1-α (ELF in the figure) promoter. ITR refers to AAV inverted terminal repeat; An refers to polyA signal; Tk refers to the HSV thymidine kinase promoter; *neo* refers to the Tn5 neomycin phosphotransferase gene.

Figure 2 shows hAAT secretion rates *in vitro* from transiently transfected murine C2C12 myoblast cell line using expression vectors according to the subject invention. C-AT does not differ significantly from E-AT, but both differ from A-AT and B-AT ($p<0.05$) AAT expression was detected using an ELISA assay specific for human AAT.

Figure 3 shows hAAT secretion rates *in vitro* from stably transduced murine C2C12 myoblast cell line using viral particles comprising expression vectors according to the subject invention. The mean rates of secretion from G418-resistant cultures 1 mo after transduction with either packaged E-AT vector or packaged C-AT vector are shown. In each instance, a "low" multiplicity transduction ($4x10^5$-particles/cell) and a high multiplicity transduction ($4x10^6$ particles/cell) were performed. E-AT "low" and "high" are greater than "high" multiplicity C-AT ($P=0.02$) but are not significantly different from each other ($n=3$). AAT expression was detected using an ELISA assay specific for human AAT.

Figure 4 shows additional constructs tested for hAAT expression. The murine myoblast C2C12 cells were grown in 35-mm wells with approximately $4x10^5$ cell per well and were transfectd with 5 μg of the appropriate plasmid DNA using Superfect transfection (Qiagen Inc., CA). Secretion of hAAT into the medium was assessed at 2 days after transfection using an antigen-capture ELISA. Each bar represents the mean of results from three experiments (triplicate in each experiment).

Data from transfection experiments indicate that the expression from p43CB-AT was at least three times higher than that from C-AT *in vitro*.

Figures 5A and 5B show sustained secretion of therapeutic levels of hAAT using either the C-AT vector or the E-AT vector in either SCID or C57BL mice. Figure 5A shows the mean total serum levels of hAAT observed in groups of either SCID (squares)

or C57BL (circles) mice receiving either low dose ($5x10^{11}$ particles) (open symbols) or high dose ($1.4x10^{13}$ particles) (filled symbols) single injections into muscle of the C-AT vector measured at time points ranging from 1 to 16 wk after injection. For each strain, the high-dose curve is significantly different from the low-dose curve (P=0.009 for SCID, P=0.02 for C57BL), but the strains do not differ from each other. Figure 5B shows analogous data with the E-AT vector. None of these differences were significant.

Figure 5C shows long term secretion of hAAT from murine muscle transduced with C-AT. C57B1/6 or C57B1/6-SCID mice received $3.5x10^{10}$ IU, $1.4x10^{13}$ particles/mouse. One year after injection, serum hAAT levels were still 400 μg/ml in C57B1/6-SCID and 200 ug/ml in C57B1/6. This level are comparable with the peak levels observed (800 or 400 ug/ml, respectively).

Figure 6 shows an immunoblot of sera taken from several of the C-AT vector-treated mice at 11 weeks after vector administration. Ten microliters of a 1:100 dilution of serum was electrophoresed by 10% SDS/PAGE, blotted, and incubated with 1:1,500 dilution of goat anti-hAAT-horseradish peroxidase conjugate (Cappel/ICN). Samples from three high-dose SCID (h1-h3), one high-dose C57BL (h3), and three low-dose C57BL (lo1-lo3) were included, along with one negative control (saline-injected = sal) serum to indicate the level of reactivity with endogenous mAAT. As a standard, hAAT was added either to negative-control C57BL serum (first hAAT lane) or to PBS (second hAAT) lane to final equivalent serum concentration of 100 μg/ml.

Figures 7A and 7B show that some BALB/c mice mount humoral immune responses to hAAT, which correlate with lower serum levels but no observable toxicity. Figure 7A shows serum hAAT levels and Figure 7B shows serum anti-hAAT antibody levels as determined by ELISA performed on serum taken from mice injected with 1 x 10^{11} particles of the C-AT vector. Each set of symbols represents an individual animal (□, no. 1; △, no. 2; ○, no. 3). Note the inverse correlation between the presence of antibody and the presence of circulating hAAT.

Figure 8 shows the persistence of rAAV-AAT vector DNA in high molecular weight form. PCR products were amplified from DNA prepared by Hirt extraction from three SCID mice injected 16 wk earlier with $5x10^{11}$ resistant-particles of C-AT and analyzed by Southern blot. The high molecular weight Hirt pellet (genomic DNA lanes) and the low molecular weight supernatant (episomal DNA lanes) were analyzed

separately. Control lanes include a sample in which an hAAT cDNA plasmid was the template DNA (+) and a control in which water was the template (-). In this internal PCR reaction, a 500-bp product is expected regardless of whether or not the vector genome is integrated.

Figure 9 shows serum hAAT in C57B1/6 mice transduced with C-AT and p43CB-AT. C57B1/6 mice were injected in muscle with C-AT ($3.5x10^{10}$ IU/mouse, $1x10^{12}$ particles/mouse) or p43CB-AT ($6x10^{9}$ IU, $1x10^{12}$ particles/mouse). The level of hAAT from p43CB-AT were projected based on an estimation of the equivalent dosage (infectious unit) of C-AT.

Figure 10 shows enhancement of CMV promoter activity by a synthetic enhancer in C2C12 cells. The murine myoblast C2C12 cells were grown in 35-mm wells with approximately $4x10^{5}$ cell per well and were transfected with 5 μg of p.43rmsENC-AT vector DNA using SUPERFECT transfection (Qiagen Inc, CA). Secretion of hAAT into the medium was assessed at 2 days after transfection using an antigen-capture ELISA. Each bar represents the mean of results from one experiment (triplicate).

Figure 11 shows secretion of hAAT from mouse liver cells (HO15) transfected with different constructs. The murine liver cells (HO15) were grown in 35-mm wells with approximately $4x10^{5}$ cell per well and were transfected with 5μg of the plasmid DNA using LIPOFECTAMINE reagents (Life Technologies Inc, MD). Secretion of hAAT into the medium was assessed at 2 days after transfection using an antigen-capture ELISA. Each bar represents the mean of results from two experiments (triplicate).

Figure 12 shows secretion of hAAT from mouse liver cells (HO15) transfected using different methods. The murine liver cells (HO15) were grown in 35-mm wells with approximately $4x10^{5}$ cell per well and were transfected with 5μg of the p43CB-AT vector using Superfect (Qiagen Inc., CA), FuGENE (Boehringer Mannhem Co, IN), Lipofectin, LipofectAMINE (Life Technologies Inc, MD) reagents and Calcium phosphate (CA-PO4) transfection. Secretion of hAAT into the medium was assessed at 2 days after transfection using an antigen-capture ELISA. Each bar represents the mean of results from one experiment (triplicate).

Figure 13 shows hAAT secretion from mouse liver transduced with rAAV. C57Bl/6 mice were injected with either p43CB-AT, C-AT or E-AT vector either by portal vein or tail vein injection. PV=portal vein injection. TV=tail vein injection.

Figure 14 shows serum hAAT levels in C57Bl/6 mice after intratracheal (IT) injection of C-AT or p43CB-AT vector. Mice received either 10^9 IU of C-AT (open circles), 10^9 IU of p43CB-AT (open triangles) or 10^{10} IU of p43CB-AT (open squares).

Figure 15 shows a map and nucleotide sequence for the vector of the present invention designated as C-AT.

Figure 16 shows a map and nucleotide sequence for the vector of the present invention designated as E-AT.

Figure 17 shows a map and nucleotide sequence for the vector of the present invention designated as dE-AT.

Figure 18 shows a map and nucleotide sequence for the vector of the present invention designated as p43C-AT.

Figure 19 shows a map and nucleotide sequence for the vector of the present invention designated as p43C-AT-IN. This vector includes intron II from human AAT gene to enhance transcription.

Figure 20 shows a map and nucleotide sequence for the vector of the present invention designated as p43CB-AT.

Figure 21 shows a map and nucleotide sequence for the vector of the present invention designated as C-AT2.

Figure 22 shows a map and nucleotide sequence for the vector of the present invention designated as p43msENC-AT. This vector is similar to p43C-AT but also comprises an enhancer sequence upstream of the CMV promoter.

Figure 23 shows a map and nucleotide sequence for the vector of the present invention designated as p43rmsENC-AT. This vector is the same as the p43msENC-AT vector except that the enhancer sequence is in an opposite orientation.

Figure 24 shows a map and nucleotide sequence for the vector of the present invention designated as p43msENCB-AT. This vector is similar to p43CB-AT but also comprises an enhancer sequence upstream of the CMV promoter.

Figure 25 shows a map and nucleotide sequence for the vector of the present invention designated as p43rmsENCB-AT. This vector is the same as p43msENCB-AT except that the enhancer sequence is in an opposite orientation.

Detailed Disclosure of the Invention

The subject invention pertains to novel materials and methods for providing gene therapy to a mammal or human having a condition or disorder, such as genetic deficiency disorders, where high levels of expression of a protein are required to treat the disorder or condition. In one method of the subject invention, a viral vector is introduced into cells of an animal wherein a therapeutic protein is produced, thereby providing genetic therapy for the animal. In one embodiment, a method of the invention comprises introducing into an animal cell or tissue an effective amount of viral particles or vector comprising a recombinant genome which includes heterologous polynucleotide encoding a protein useful in genetic therapy and that can be expressed by the cell or tissue. Expression of the heterologous polynucleotide results in production of the protein. Preferably, the therapeutic protein encoded by the heterologous polynucleotide is a serum protein. In a preferred embodiment, vector material comprising the heterologous polynucleotide is integrated into a chromosome of the cell of the host animal.

In one embodiment, a recombinant polynucleotide vector of the present invention is derived from adeno-associated virus (AAV) and comprises a constitutive or regulatable promoter capable of driving sufficient levels of expression of the heterologous DNA in the viral vector. Preferably, a recombinant vector of the invention comprises inverted terminal repeat sequences of AAV, such as those described in WO 93/24641. In a preferred embodiment, a vector of the present invention comprises polynucleotide sequences of the pTR-UF5 plasmid. The pTR-UF5 plasmid is a modified version of the pTR_{BS}-UF/UF1/UF2/UFB series of plasmids (Zolotukhin *et al.*, 1996). The pTR-UF5 plasmid contains modifications to the sequence encoding the green fluorescent protein (GFP).

Promoters useful with the subject invention include, for example, the cytomegalovirus immediate early promoter (CMV), the human elongation factor 1-alpha promoter (EF1), the small nuclear RNA promoters (Ula and Ulb), α-myosin heavy chain promoter, Simian virus 40 promoter (SV40), Rous sarcoma virus promoter (RSV), adenovirus major late promoter, β-actin promoter and hybrid regulatory element comprising a CMV enhancer/β-actin promoter. These promoters have been shown to be active in a wide range of mammalian cells. In addition to the natural promoters described above, synthetic promoters can be used in the present invention. For example, a synthetic

enhancer randomly assembled from Spc5-12-derived elements including muscle-specific elements, serum response factor binding element (SRE), myocyte-specific enhancer factor-1 (MEF-1), myocyte-specific enhancer factor -2 (MEF-2), transcription enhancer factor-1 (TEF-1) and SP-1 (Li et al., 1999; Deshpande *et al.*, 1997; Stewart *et al.*, 1996; Mitchell *et al.*, 1989; Briggs et al., 1986; Pitluk *et al.*, 1991) can be used in vectors of the invention.

The promoters are operably linked with heterologous DNA encoding the protein of interest. By "operably linked," it is intended that the promoter element is positioned relative to the coding sequence to be capable of effecting expression of the coding sequence.

Promoters particularly useful for expression of a protein in muscle cells include, for example, hybrid CMV/β-actin promoters, CMV promoters, synthetic promoters and EF1 promoter. Promoters particularly useful for expression of a protein in liver cells include, for example, hybrid CMV/β-actin promoters and EF1 promoters.

Also contemplated for use with the vectors of the present invention are inducible and cell type specific promoters. For example, Tet-inducible promoters (Clontech, Palo Alto, CA) and VP16-LexA promoters (Nettelbeck *et al.*, 1998) can be used in the present invention.

The vectors can also include introns inserted into the polynucleotide sequence of the vector as a means for increasing expression of heterologous DNA encoding a protein of interest. For example, an intron can be inserted between a promoter sequence and the region coding for the protein of interest on the vector. Introns can also be inserted in the coding regions. Transcriptional enhancer elements which can function to increase levels of transcription from a given promoter can also be included in the vectors of the invention. Enhancers can generally be placed in either orientation, 3' or 5', with respect to promoter sequences.

Heterologous polynucleotide in the recombinant vector can include, for example, polynucleotides encoding normal, functional proteins which provide therapeutic replacement for normal biological function in animals afflicted with genetic disorders which cause the animal to produce a defective protein, or abnormal or deficient levels of that protein. Proteins, and the polynucleotide sequences that encode them, which can be provided by gene therapy using the subject invention include, but are not limited to, anti-

proteases, enzymes, structural proteins, coagulase factors, interleukins, cytokines, growth factors, interferons, and lymphokines. In an exemplified embodiment, heterologous DNA in a recombinant AAV vector encodes human alpha-1-antitrypsin protein.

The gene therapy methods of the invention can be performed by *ex vivo* or *in vivo* treatment of the patient's cells or tissues. Cells and tissues contemplated within the scope of the invention include, for example, muscle, liver, lung, skin and other cells and tissues that are capable of producing and secreting serum proteins. The vectors of the invention can be introduced into suitable cells, cell lines or tissue using methods known in the art. The viral particles and vectors can be introduced into cells or tissue *in vitro* or *in vivo*. Methods contemplated include transfection, transduction, injection and inhalation. For example, vectors can be introduced into cells using liposomes containing the subject vectors, by direct transfection with vectors alone, electroporation or by particle bombardment. In an exemplified embodiment, muscle cells are infected *in vivo* by injection of viral particles comprising recombinant vector into muscle tissue of an animal. In another embodiment, liver cells are infected *in vivo* by injection of recombinant virus into either the portal vein or peripheral veins.

The methods and materials of the subject invention can be used to provide genetic therapy for any conditions or diseases treatable by protein or cytokine infusion such as, for example, alpha-1-antitrypsin deficiency, hemophilia, adenosine deaminase deficiency, and diabetes. The methods and materials of the subject invention can also be used to provide genetic therapy for treating conditions such as, for example, cancer, autoimmune diseases, neurological disorders, immunodeficiency diseases, and bacterial and viral infections. For example, the present invention can be used to provide genetic therapy to a patient wherein cells from the patient are transformed to express and produce interleukins such as interleukin-2.

Animals that can be treated with the materials and methods of the invention include mammals such as bovine, porcine, equine, ovine, feline and canine mammals. Preferably, the mammals are primates such as chimpanzees and humans.

The subject invention also concerns cells containing recombinant vectors of the present invention. The cells can be, for example, animal cells such as mammalian cells. Preferably, the cells are human cells. More preferably, the cells are human myofibers or myoblasts, hepatocytes or lung cells. In a preferred embodiment, a recombinant vector

of the present invention is stably integrated into the host cell genome. Cell lines containing the recombinant vectors are also within the scope of the invention.

In an exemplified embodiment, recombinant AAV vectors comprising the human AAT gene (hAAT) using either the CMV promoter (AAV-C-AT) or the human elongation factor 1-alpha (EF1) promoter (AAV-E-AT) to drive expression were constructed and packaged using standard techniques. A murine myoblast cell line, C2C12, was transduced with each vector and expression of hÀAT into the medium was measured by ELISA. *In vitro*, the EF1 promoter construct resulted in 10-fold higher hAAT expression than the CMV promoter construct. *In vivo* transduction was performed by injecting doses of up to 1.4×10^{13} Dnase-resistant particles of each vector into skeletal muscles of a number of different strains of mice (including C57B1/6, Balb/c, and SCID). *In vivo*, the CMV promoter construct resulted in higher levels of expression, with sustained serum levels up to 800 μg/ml in SCID mice, approximately 10,000-fold higher than those previously observed with proteins secreted from AAV vectors in muscle. At lower doses in both C57B1/6 and SCID mice, expression was delayed for several weeks, but was sustained for over 10 weeks without declining. Thus, increasing dosage AAV vector via transduction of skeletal muscle provides a means for replacing AAT or other serum proteins.

Transduction of muscle using the vectors of the subject invention presents several advantages in that it is stable, non-toxic, and relatively nonimmunogenic. Furthermore, certain transcription promoters, such as the CMV promoter, which appear to be markedly down-regulated in other contexts have been found to remain active over time as used in the subject invention. Using the materials and methods of the subject invention, microgram/ml serum levels of a therapeutic protein can be achieved. In an exemplified embodiment, the levels of *in vivo* protein expression achieved represent a 10,000-fold or more increase over previously published results. In addition, a dose-effect relationship was demonstrable within the range of doses used, providing for further increases in expression levels as vector dose is increased.

In another embodiment of the invention, recombinant AAV vectors *i.e.*, C-AT, p43C-AT, P43CB-AT, E-AT and dE-AT comprising the human AAT gene (hAAT) using were constructed and packaged using standard techniques. A murine liver cell line, HO15, was transfected with each vector and expression of hÀAT into the medium was

measured by ELISA. *In vitro*, transduction with the p43CB-AT vector exhibited the highest level of hAAT expression. *In vivo*, the p43CB-AT vector also gave higher levels of expression. Portal vein administration appeared to be the more efficient route of administration as mice injected in this manner exhibited higher levels of expression than those receiving peripheral vein injections. Transduction of liver offers the same advantages as for muscle, but hepatocytes may be more efficient at secretion of protein.

The dosage of recombinant vector or the virus to be administered to an animal in need of such treatment can be determined by the ordinarily skilled clinician based on various parameters such as mode of administration, duration of treatment, the disease state or condition involved, and the like. Typically, recombinant virus of the invention is administered in doses between 10^5 and 10^{14} infectious units. The recombinant vectors and virus of the present invention can be prepared in formulations using methods and materials known in the art. Numerous formulations can be found in Remington's Pharmaceutical Sciences, 15th Edition (1975).

All publications and patents cited herein are expressly incorporated by reference.

Materials and Methods

Construction of rAAV plasmids. The rAAV-AAT vector plasmids used for these experiments are depicted diagrammatically (Figure 1). Briefly, the plasmid pN2FAT (Garver *et al.* (1987) plasmid was digested with *Xho*I to release 1.8-kb fragment containing the human AAT cDNA along with the SV40 promoter and a polyadenylation signal. This fragment was subcloned into a plasmid, pBlueScript (Stratagene) and, after the removal of the SV40 promoter by *Hind* III digestion and religation, the hAAT cDNA with its polyA signal was released by *Xba*I and *Xho*I digestion. This 1.4-kb *Xba*I-*Xho*I fragment was then cloned in to the pTR-UF5 (an AAV-inverted terminal repeat-containing vector) plasmid (Zolotukhin *et al.*, 1996) between the *Xba*I site 3' to the CMV promoter and the *Xho*I site 5' to the polyoma virus enhancer/HSVthymidine kinase promoter cassette, which drives *neo* in that construct. This yielded the pAAV-CMV-AAT construct (C-AT). Analogous constructs using the promoter from the small nuclear RNA proteins, Ula and Ulb, (to give the A-AT and B-AT constructs, respectively) and

human elongation factor 1-alpha (EF1) promoter (to give the E-AT construct) were constructed by substituting each of these promoter cassettes in place of the CMV promoter, between the KpnI and XbaI sites.

The construct, dE-AT derived from E-AT by deletion of the silencer (352 bp) by SAC II-cut (Wakabayashi-Ito *et al.*, 1994). C-AT2 is similar with C-AT except there are SV40 intron and poly (A) sequences flanking the cDNA of hAAT. The p43C-AT was constructed by insertion of hAAT cDNA to an AAV-vector plasmid (p43), which has CMV promoter, intron and poly (A) sequences. The p43CB-AT is derived by replacement of CMV promoter with CMV enhancer and chicken β-actin promoter sequences. The p43C-AT-IN is derived from p43C-AT by insertion of intron II sequences of hAAT gene to hAAT cDNA (Brantly *et al.*, 1995).

Packaging of rAAV vectors. Vectors were packaged using a modification of the method described by Ferrari *et al.* (1997) . Briefly, plasmids containing the AAV *rep* and *cap* genes (Li *et al.*, 1997) and the Ad genes (E2a, E4 and VA-RNA) were co-transfected along with the appropriate AAV-AAT vector plasmid into 293 cells grown in Cell Factories (Nunc). Cells were harvested by trypsinization and disrupted by freeze-thaw lysis to release vector virions which were then purified by iodixanol gradient ultracentrifugation followed by heparin sepharose affinity column purification. Alternatively, recombinant virus can be prepared according to methods described in Zolotukhin *et al.* (1999).

Vector preparations had their physical titer assessed by quantitative competitive PCR and their biological titer assessed by infectious center assay. The presence of wild-type AAV was also assessed using these same assays with appropriate internal AAV probes. The high-dose C-AT stock had a particle-titer of $2.0x10^{14}$ particles/ml and an infectious titer of $5.0x10^{11}$ infectious units (i.u.)/ml (particle to i.u. ratio = 400:1). The low-dose C-AT measured $8x10^{12}$ particles/ml and $1.2x10^{10}$ i.u./ml (particle to i.u. = 667:1). For the E-AT experiments, the titers were $1x10^{13}$ particles/ml and $2.5x10^{10}$ i.u./ml (particle to i.u. = 400:1). The low-dose C-AT stock had a wt-like AAV particle titer (*i.e.*, positive AAV genome PCR) equal to 0.1 times the recombinant titer but no detectable infectious wtAAV. The other two preparations had wt-like AAV particle titers $<10^{-5}$ times the recombinant titer and no detectable infectious wtAAV.

In vitro transfection and transduction experiments. The C2C12 murine myoblast line was used for *in vitro* transfection and transduction experiments. Cells were grown in 35-mm wells with approximately 4 x 10^5 cells per well and transfected with 5 μg of each plasmid DNA using Superfect (Qiagen Corp.). Secretion of hAAT into the medium was assessed at 2 days after transfection using an antigen-capture ELISA assay with standards (Brantly *et al.*, 1991). An SV40 promoter luciferase-expression plasmid, pGL2 (Promega), was used as an internal control. For transduction experiments, cells were grown under similar conditions and were transduced with vector at multiplicities of infection ranging from 4 x 10^5 to 4 x 10^6 particles per cell. Cells were then passaged in the presence of geneticin sulfate (350 μg/ml) and geneticin-resistant clones were isolated for hAAT secretion studies.

In vivo injection of AAV-C-AT and AAV-E-AT vectors into murine muscle. Mice strains (C57B1/6, SCID, and Balb/c) were obtained from Jackson Laboratories (Bar Harbor, ME) and were handled under specific pathogen-free conditions under a protocol approved by the University of Florida Institutional Animal Care and Use Committee. Animals were anesthetized by metaphane inhalation and aliquots of vector were injected percutaneously into the quadriceps femoris muscles of both hind limbs. The volume of vector ranged from 50 to 100 μl per injection site and the total amount of virus injected per animal ranged from 5 x 10^{10} to 1.4 x 10^{13} Dnase-resistant particles.

Antigen capture ELISA assay for hAAT expression. Microtiter plates (Immulon 4, Dynex Technologies, Chantilly, VA) were coated with 100 μl of a 1:200 dilution of goat anti-human AAT (CAPPEL/ICN) in Vollers buffer (Na2CO3=2.76g, NaHCO3=1.916g, NaN3=0.2g, d.H2O=1 liter, Adjust PH=9.6) overnight at 4°C. After washing, standards and unknown samples containing hAAT were incubated in the plates at 37°C for 1 hour. After blocking in 3% BSA in PBS-Tween 20 at 37°C for 1 hour, a second antibody (1:1000 dilution of rabbit anti-human AAT, Boehringer Mannheim) was reacted with the captured antigen at 37°C for 1 hour. Detection was performed using a third antibody incubation (1:800 dilution of goat anti-rabbit IgG-peroxidase conjugate, 37°C) followed by *o*-phenylenediamine (OPD, Sigma) detection and measurement of the absorbance at 490nm.

ELISA assay for anti-hAAT and anti-AAV VP3 antibodies. Wells were coated with antigen (1 μg of hAAT or 100ng of VP3) at 4°C overnight, blocked with 3% BSA

and then reacted with dilutions of either test serum or with positive control antibodies at 37°C for 1 hour. After washing, a goat-anti-mouse IgG-peroxidase conjugate was used as a secondary antibody (1:1500 dilution) to detect bound anti-AAT antibody, using a standard OPD reaction, as described above. Antibody levels were quantitated by comparison with a standard curve generated by reacting dilutions of known positive monoclonal antibodies against VP3 and hAAT.

Lymphocyte proliferation assays to detect cell-mediated immune responses. Lymphocyte proliferation assays were performed in order to detect T cell responses to the hAAT and VP3 antigens. Freshly isolated splenocytes were grown in primary culture in 96 well plates coated with 0, 0.1, 1, and 10 μg of either hAAT or VP3 in RPMI-C+ medium. On day three, a pulse of ^{3}H-thymidine was added, and the cells were harvested on day 4 for lysis and scintillation counting. Phytohemagglutinin (PHA) was used as a mitogen for positive control wells. A stimulation index was calculated for each antigen dosage level by dividing the counts per minute (cpm) of ^{3}H-thymidine incorporated in the antigen-stimulated cells by the cpm in a control (unstimulated) well.

Following are examples which illustrate procedures for practicing the invention. These examples should not be construed as limiting. All percentages are by weight and all solvent mixture proportions are by volume unless otherwise noted.

Example 1 – *In vitro* studies in murine C2C12 myoblasts

In order to determine the relative strength of a number of constitutively active promoters in the context of AAV-AAT vectors, packageable AAV-AAT expression vectors containing one of the CMV, EF1, Ula or Ulb promoters (Figure 1) were constructed. Each of these constructs were transfected in to the murine C2C12 myoblast cell line. Both the EF1 and the CMV promoter were active for AAT expression, with EF1 construct (AAV-E-AT) expressing 850 ng/10^5 cells/day and the CMV construct (AAV-C-AT) expressing approximately 670 ng/10^5 cells/day, as measured by a human-specific ELISA assay for AAT (Figure 2). This difference was not statistically significant. The levels of expression from the Ula and Ulb constructs were undetectable.

In order to better characterize the level and duration of expression in the setting of vector transduction, cultures of C2C12 cells were transduced with either AAV-E-AT

or AAV-C-AT at multiplicities of infection ranging from $4x10^5$ to 4×10^6 Dnase-resistant particles per cell. Cells were then selected for expression of the *neo* gene (present in each of the AAV constructs) by growth in G418-containing medium. Several cell clones and pooled cell populations were independently analyzed for AAT expression at four weeks post-transduction (Figure 3). There was a clear trend toward higher levels of expression at higher multiplicities of infection, and the E-AT construct expressed at least 10-fold greater quantities under all conditions in these long-term cultures. The most active E-AT clone expressed hAAT at a rate of over 1400 ng/10^5 cells/day.

Example 2— *In vivo* expression of hAAT from murine skeletal muscle

In order to determine whether the AAV-AAT constructs would be active *in vivo* in skeletal muscle, doses of vector were injected into the quadriceps femoris muscle of mice. Circulating serum levels of hAAT were then measured for 11 to 15 weeks after the initial injection. Four saline-injected animals from each mouse strain served as controls. In the case of the C-AT vector (Figure 5A), levels of expression were sufficient to achieve serum levels in excess of 800 μg/ml in SCID mice after a single injection of 1.4×10^{13} particles. A dose-effect relationship was observed, with expression levels in SCID being at least 20-fold lower at the 5×10^{11} particle dose. The levels of expression increased over the first several weeks after injection and were stable thereafter until the time of sacrifice. Since hAAT has a half-life of less than 1 week, this indicated continuous expression. Levels from C57B1/6 mice were comparable, and also achieved values close to the therapeutic range. In similar studies, two of three Balb/c mice injected with 1×10^{11} particles of the C-AT vector did not express hAAT at detectable levels. Both of these were found to have developed high levels of anti-hAAT antibodies.

Surprisingly, expression levels from the AAV-E-AT vector after *in vivo* injection were modestly lower than those seen with the C-AT vector (Figure 5B), with maximal levels of approximately 250 ng/ml at the 5×10^{11} dose at and beyond 7 weeks in SCID mice. When the dose was further increased to 1×10^{12} particles, levels of approximately 1200 ng/ml were observed. These levels were stable for one year post-injection (Figure 5C). Levels observed in SCID and immune competent C57B1/6 mice were similar.

Example 3 — Immunologic Studies

In studies in Balb/c mice, antibody levels against hAAT were high in 2 of 3 animals injected. The one which did not have circulating anti-hAAT was the only animal with levels of hAAT expression similar to those in the C57B1/6 and SCID groups. The high-dose C57-C-AT injection group had detectable levels of antibody directed against VP3, but not hAAT.

In order to determine whether any cell-mediated immune responses were mounted, lymphocyte proliferation assays were performed using either hAAT or AAV-VP3 for antigenic stimulation of primary splenic lymphocytes harvested at the time of animal sacrifice, 16 weeks post-vector injection. Using this method, no immune responses were detectable in any of the mice.

Example 4 — Lack of toxicity from direct vector injection

In order to determine whether there was any direct toxicity, inflammation, or neoplastic change associated with vector injection, animals underwent complete necropsies. Histopathologic examination was performed on 5 μm sections taken from the site of vector injection and from a panel of other organs, including the brain, heart, lungs, trachea, pancreas, spleen, liver, kidney, and jejunum. No histologic abnormalities were observed in any of these sites, even among those mice which developed humanol immune responses against hAAT.

Example 5 — Molecular evidence of AAV-AAT vector persistence

To confirm the presence of vector DNA, a vector-specific PCR (*neo* primers 5'-TATGGGATCGGCCATTGAAC-3', and 5'-CCTGATGCTCTTC-GTCCAGA-3', was performed on DNA extracted from 3 SCID mice 16 weeks after injection with the C-AT vector, and PCR products were analyzed by Southern blot analysis with a ^{32}P-labeled vector-specific probe (Figure 8). The state of vector DNA was analyzed using the Hirt procedure (Carter *et al*, 1983) to separate the low molecular weight episomal DNA from the high molecular weight fraction, which would contain integrated forms and large concatemers. In each case, vector DNA was present in the high molecular weight DNA fraction, whereas in only one of the animals was there a signal in the episomal fraction. This result indicates that by 16 weeks most of the vector DNA in our animals was either integrated or in large concatemers.

Example 6 — *In vivo* expression of hAAT from murine liver

Portal vein or tail vein injections were performed on 18 female C57BL/6 mice 8-10 weeks of age. The injection volume was 100μl per mouse.

Each group had the following parameters:

1. Group 1: 100 μl of PBS n=4.
2. Group 2: 100 μl of p43CB-AT ($3x10^{10}$ IU/animal) n=3.
3. Group 3: 100 μl of p43CB-AT ($4x10^9$ IU/animal) n=4.
4. Group 4: 100 μl of C-AT ($4x10^9$ IU/animal) n=2.
5. Group 5: 100 μl of E-AT ($4x10^9$ IU/animal) n=4.
6. Group 6: EATM TV=100 μl by tail vein injection of E-AT ($4x10^9$ IU/animal) n=3.
7. Group 0: 100 μl of PBS by tail vein injection n=2.

A total of 22 animals were used in this study.

All animals were anesthetized with 2-2-2 tribromoethanol (Avertin) using a working solution of 20 mg/ml at a dosage of 0.5 mg/g IP. A 2 cm ventral midline abdominal incision was made from the pubic symphysis extending cranially to the xyphoid process through skin and muscle layers. The portal vein was exposed by retracting the intestines and associated mesentery to the left side of the animal. Additionally, the quadrate and right medial lobes of the liver were retracted cranially. Intestines and peritoneal cavity were continuously lavaged with 0.9% NaC.1

Virus or PBS was delivered into the portal vein using a 30 g needle attached to a 100 ul capillary pipette using mouth delivery via rubber tubing and a Drummond self-locking double layer 0.8 um filter. A small piece of Gel-Foam (.5x.5cm) was applied to the injection site before the needle was removed from the portal vein. The needle was retracted from beneath the Gel-Foam and the piece was held in place with forceps while the intestines were replaced into the peritoneal cavity.

The muscle and skin were closed in one layer using 2 simple interrupted 3-0 nylon sutures on an FS-1 cutting needle. Surgeries were performed on a thermoregulated operating board designed to maintain a temperature of 37 degrees. For recovery from anesthesia, the animals were placed under a heat lamp adjusted to maintain an ambient temperature of approximately 37 degrees and given subcutaneous fluid if there was a significant amount of blood loss during surgery.

Serum levels of hAAT in the mice were measured two weeks after injection. Serum levels of about 200-150 μg/ml hAAT were detected in mice receiving the p43CB-AT vector (Figure 13). Studies using the E-AT vector show that injection of vector by portal vein led to greater levels of hAAT secretion as compared to E-AT administered by tail vein injection.

Example 7— *In vivo* expression of hAAT from murine lung

Mice were injected intratracheally with either C-AT or p43CB-AT vector. Serum levels of hAAT in the mice were measured at day 3, 14 and 31 after injection (Figure 14). The p43CB-AT vector mediated high levels of expression of hAAT in lung.

It should be understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and the scope of the appended claims.

References

U.S. Patent No. 5,399,346

U.S. Patent No. 5,439,824

U.S. Patent No. 5,658,785

U.S. Patent No. 5,858,351

Afione, S. A., Conrad, C. K., Kearns, W. G., Chunduru, S., Adams, R., Reynolds, T. C., Guggino, W. B., Cutting, G. R., Carter, B. J. and Flotte, T. R. (1996) *J. Virol.* 70:3235-3241.

Brantly, M. L., Wittes, J. T., Vogelmeier, C. F., Hubbard, R. C., Fells, G. A., Crystal, R. G. (1991) *Chest* 100:703-708.

Briggs, M. R., Kadonga, J. T., Bell, S. P., Tjian, R. (1986) *Science* 234:47-52.

Carter *et al.* (1983) *Virology* 126:505-515.

Deshpande, N., Chopra, A., Rangarajan, A., Shashidhara, L. S., Rodrigues, V., Krishna, S. (1997) *J. Biol. Chem* 272(16):10664-10668.

Garver, R. I., Jr., Chytil, A., Courtney, M., Crystal, R. G. (1987) *Science* 237:762-764.

Ferrari, F. K., Samulski, T., Shenk, T., Samulski, R. J. (1996) *J. Virol.* 70:3227-3234.

Ferrari, F. K., Xiao, X., McCarty, D. M., Samulski, R. J. (1997) *Nature Med* 3:1295-97.

Fisher, K. J., Gao, G. P., Weitzman, M. D., DeMatteo, R., Burda, J. F., Wilson, J. M. (1996) *J. Virol.* 70:520-532.

Flotte, T. R., Afione, S. A., Conrad, C., McGrath, S. A., Solow, R., Oka, H., Zeitlin, P. L., Guggino, W. B., Carter, B. J. (1993) *Proc. Natl. Acad. Sci. USA* 90:10613-10617.

Jooss, K., Yang, Y., Fisher, K. J., Wilson, J. M. (1998) *J. Virol.* 72:4212-4223.

Li, X, Eastman, E. M., Schwartz, R. J., Draghia-Akli, R. (1999) *Nat. Biotechnol* 17(3):241-245.

Li *et* al. (1997) *J. Virol.* 71:5236-43.

Kearns, W. G., Afione, S. A., Fulmer, S. B., Caruso, J., Flotte, T. R., Cutting, G. R. (1996) *Gene Ther.* 3:748-755.

Kessler, P. D., Podsakoff, G., Chen, X., McQuiston, S. A., Colosim, P. C., Matelis, L. A., Kurtzman, G. and Byrne, B. J. (1996) *Proc. Natl. Acad. Sci. USA* 93:14082-14087.

Klein, R. L. *et al.* (1998) *Exp. Neurol.* 150:183-194.

Kotin, R. M., Siniscalco, M., Samulski, R. J., Zhu, X. D., Hunter, L., Laughlin, C. A., McLaughlin, S., Muzyczka, N., Rocchi, M., Berns, K. I. (1990) *Proc. Natl. Acad. Sci. USA* 87:2211-2215.

Mitchell, P. J., Tjian, R. (1989) *Science* 245:371-378.

Murphy, J. E., Zhou, S., Giese, K., Williams, L. T., Escobedo, J. A., Dwarki, V. J. (1997) *Proc. Natl. Acad. Sci. USA* 94:13921-13926, 1997.

Nettelbeck, D. M., Jerome V., Muller R. (1998) *Gene Ther* 5(12)1656-1664.

Pitluk, Z. W., Ward, D. C. (1991) *J. Virol.* 65:6661-6670.

Stewart, A. F., Richard, III, C. W., Suzow, J., Stephan D., Weremowicz, S., Morton, C. C., Andra, C. N. (1996) *Genomics* 37(1):68-76.

Xiao, X., Li, J., Samulski, R. J. (1996) *J. Virol.* 70:8098-8108.

Zolotukhin, S., Potter, M., Hauswirth, W., Guy, J., Muzyczka, N. (1996) *J. Virol.* 70(7):4646-4654.

Zolotukhin, S. *et al.* (1999) *Gene Ther.* (6).

We claim:

Claims

1. A method for providing an animal with a therapeutically effective amount of a serum protein, said method comprising introducing into cells of said animal an effective amount of viral particles or vector, wherein said viral particles or viral vector comprises a polynucleotide encoding said protein.

2. The method according to claim 1, wherein said animal is a mammal.

3. The method according to claim 2, wherein said mammal is a human.

4. The method according to claim 1, wherein said vector is an adeno-associated virus vector.

5. The method according to claim 1, wherein said vector comprises a promoter sequence capable of driving expression of said polynucleotide encoding said protein.

6. The method according to claim 5, wherein said promoter sequence is selected from the group consisting of CMV promoter sequences, hybrid CMV enhancer/β-actin promoter sequences, EF1 promoter sequences, Ula promoter sequences and U1b promoter sequences.

7. The method according to claim 5, wherein said promoter sequence is an inducible promoter selected from the group consisting of Tet-inducible promoters and VP16-LexA promoters.

8. The method according to claim 5, wherein said vector further comprises an enhancer sequence.

9. The method according to claim 8, wherein said enhancer is a synthetic enhancer.

10. The method according to claim 1, wherein said animal has a condition that results in a defective protein or a deficiency of said protein encoded by said polynucleotide.

11. The method according to claim 1, wherein said animal has a condition that can be ameliorated or treated by said protein encoded by said polynucleotide.

12. The method according to claim 1, wherein said protein encoded by said polynucleotide is selected from the group consisting of anti-proteases, enzymes, structural proteins, coagulase factors, interleukins, cytokines, growth factors, interferons, and lymphokines.

13. The method according to claim 1, wherein said cells are myofibers, myoblasts, hepatocytes, or lung cells.

14. The method according to claim 1, wherein said polynucleotide encodes human alpha-1-antitrypsin protein, or a biologically active fragment or variant thereof.

15. The method according to claim 4, wherein said polynucleotide encodes human alpha-1-antitrypsin protein, or a biologically active fragment or variant thereof.

16. The method according to claim 1, wherein said viral particles are introduced into said cells or tissue by infection or injection.

17. The method according to claim 1, wherein said vector is introduced into said cells by transfection or injection.

18. The method according to claim 1, wherein said viral particles or vector is introduced into said cells *in vitro* and said treated cells are introduced into said animal.

19. The method according to claim 1, wherein said viral particles or vector is introduced into said cells *in vivo*.

20. The method according to claim 19, wherein said viral particles or vector is injected into muscle.

21. The method according to claim 19, wherein said viral particles or vector is injected into portal or peripheral vein.

22. The method according to claim 19, wherein said viral particles or vector is injected intratracheally or inhaled into the lungs.

23. The method according to claim 15, wherein said vector is selected from the group consisting of dE-AT, E-AT, C-AT, C-AT2, p43C-AT, p43CB-AT, p43C-AT-IN, p43msENC-AT, p43rmsENC-AT, p43msENCB-AT and p43rmsENCB-AT.

24. A recombinant viral vector comprising a polynucleotide encoding a protein capable of providing a therapeutic effect to an animal when expressed in said animal.

25. The vector according to claim 24, wherein said animal is a mammal.

26. The vector according to claim 25, wherein said mammal is a human.

27. The vector according to claim 26, wherein said vector is an adeno-associated virus vector.

28. The vector according to claim 24, wherein said vector comprises a promoter sequence capable of driving expression of said polynucleotide encoding said protein.

29. The vector according to claim 24, wherein said promoter sequence is selected from the group consisting of CMV promoter sequences, hybrid CMV enhancer/β-actin

promoter sequences, EF1 promoter sequences, Ula promoter sequences and U1b promoter sequences.

30. The vector according to claim 24, wherein said polynucleotide encodes human alpha-1-antitrypsin protein, of a biologically active fragment or variant thereof.

31. The vector according to claim 27, wherein said polynucleotide encodes human alpha-1-antitrypsin protein, of a biologically active fragment or variant thereof.

32. The vector according to claim 31, wherein said vector is selected from the group consisting of dE-AT, E-AT, C-AT, C-AT2, p43C-AT, p43CB-AT, p43C-AT-IN, p43msENC-AT, p43rmsENC-AT, p43msENCB-AT and p43rmsENCB-AT.

33. A viral particle comprising the vector of claim 24.

34. A cell comprising the vector of claim 24.

35. The cell according to claim 34, wherein said cell is a myofiber, myoblast, hepatocyte, or lung cell.

36. A method for treating alpha-1-antitrypsin deficiency in an animal, said method comprising introducing into cells of said animal a vector according to claim 24, wherein said polynucleotide of said vector encodes alpha-1-antitrypsin protein, or a biologically active fragment or variant thereof.

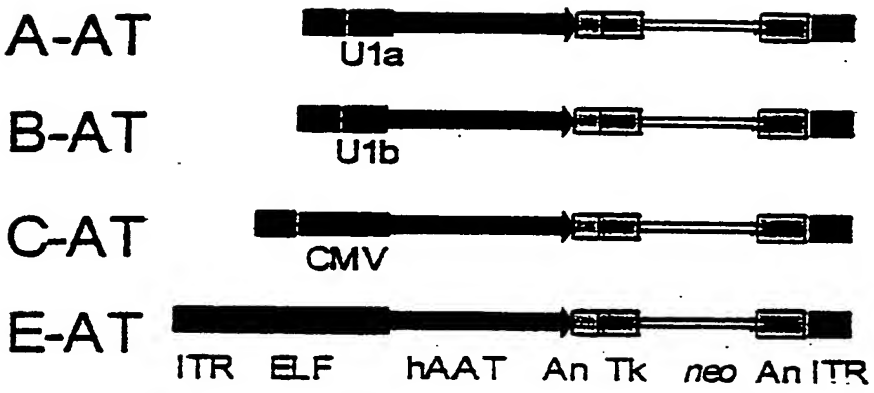


FIGURE 1

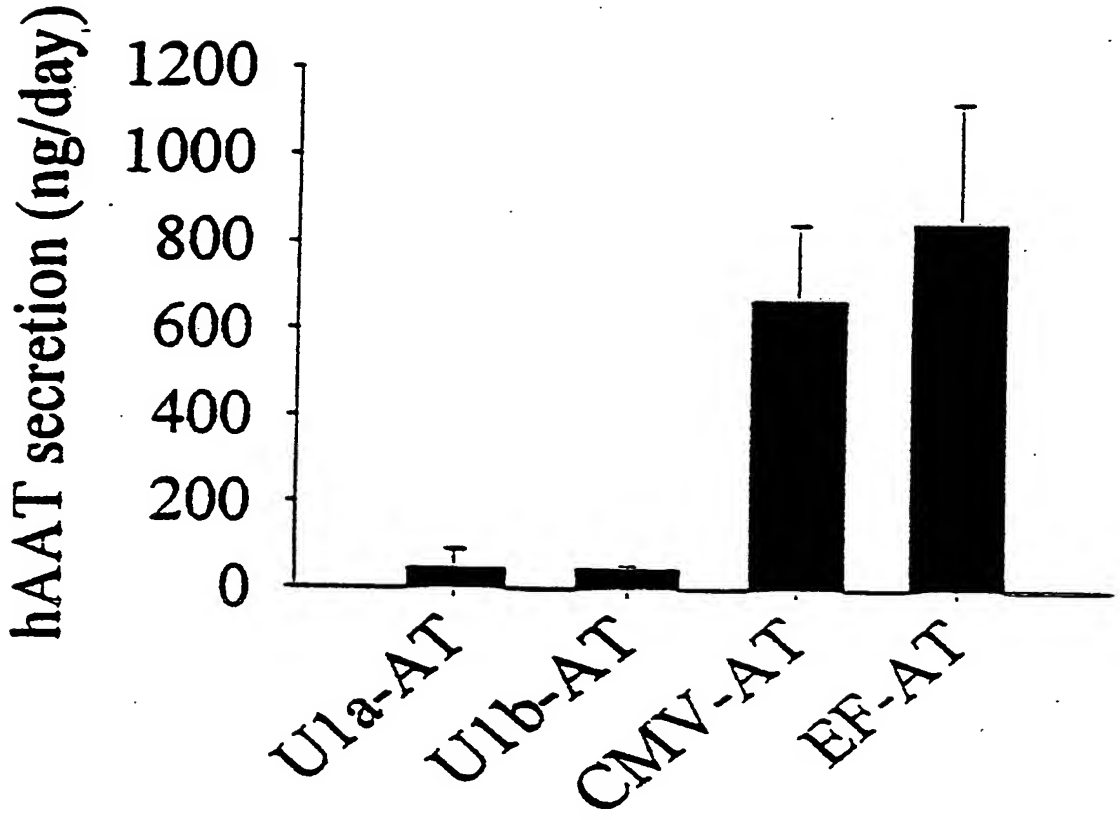

hAAT secretion (ng/day)
1200
1000
800
600
400
200
0
U1a-AT
U1b-AT
CMV-AT
EF-AT

FIGURE 2

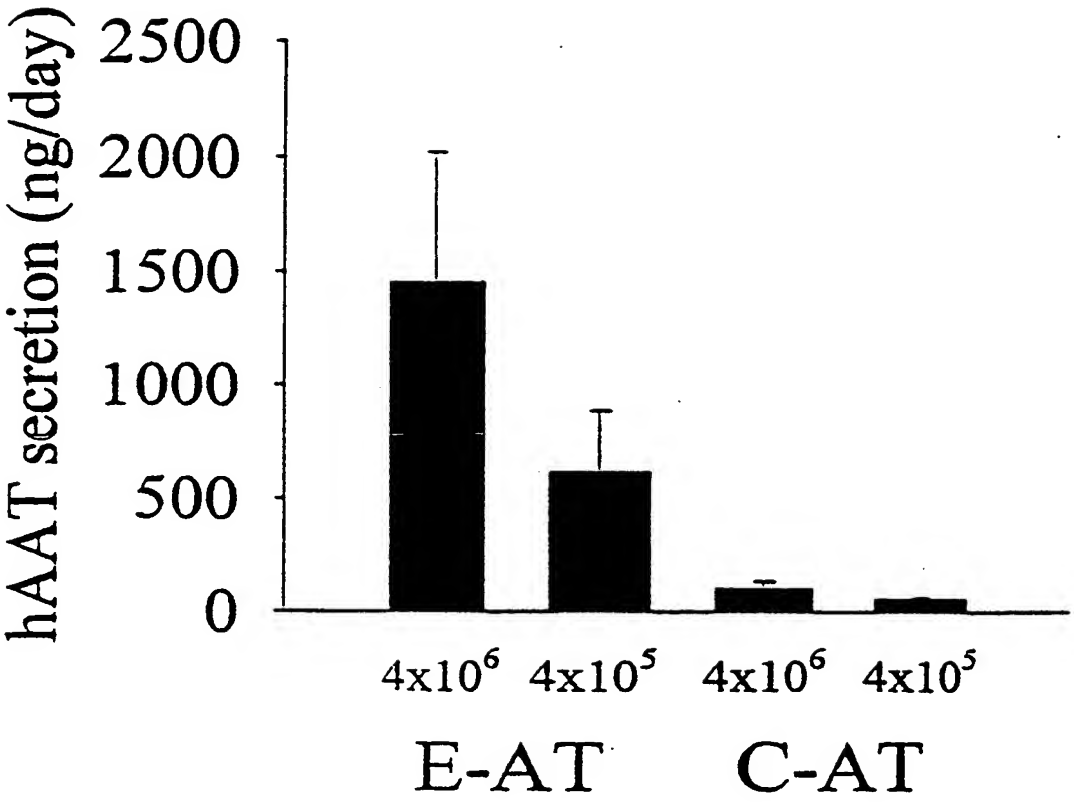

hAAT secretion (ng/day)
2500
2000
1500
1000
500
0
4x10^6
4x10^5
4x10^6
4x10^5
E-AT
C-AT

FIGURE 3

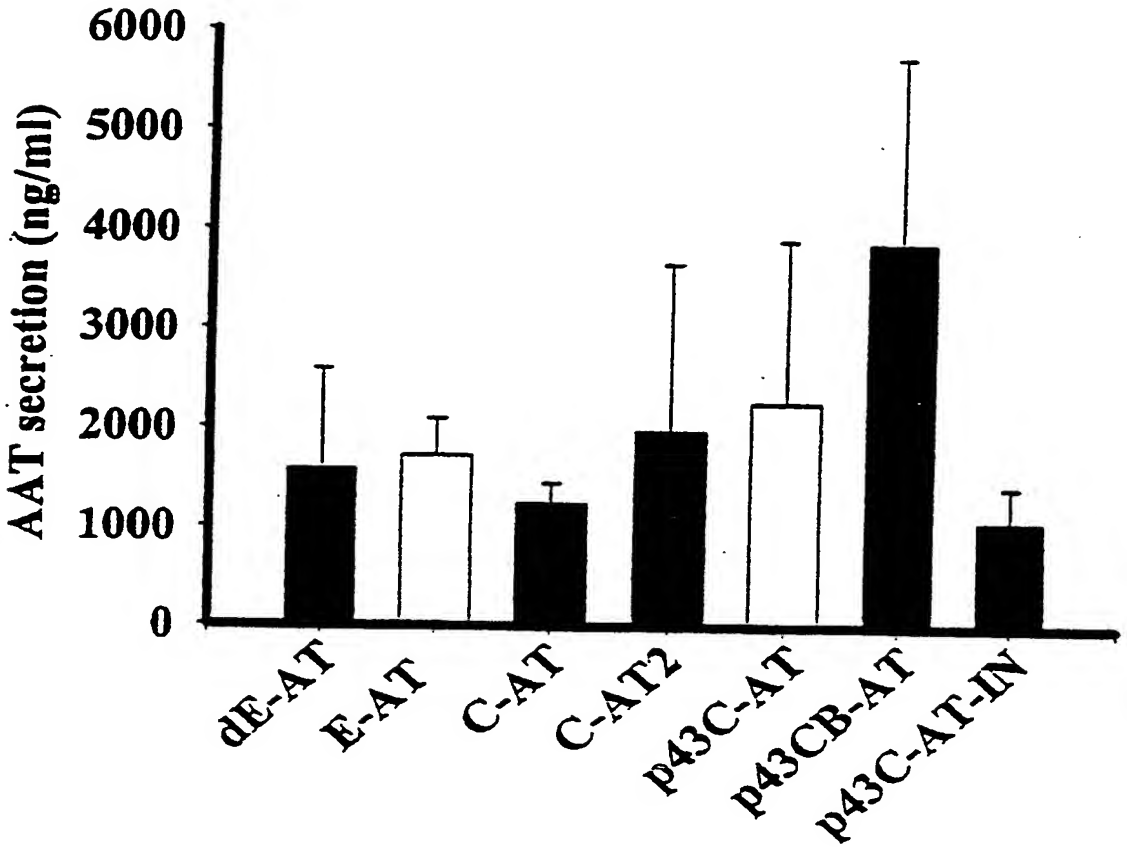

AAT secretion (ng/ml)
6000
5000
4000
3000
2000
1000
0
dE-AT
E-AT
C-AT
C-AT2
p43C-AT
p43CB-AT
p43C-AT-IN

FIGURE 4

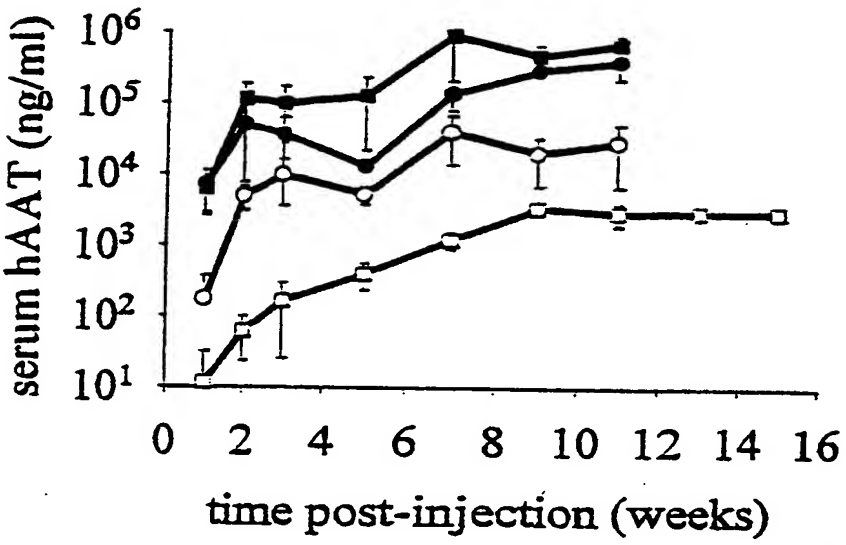


FIGURE 5A

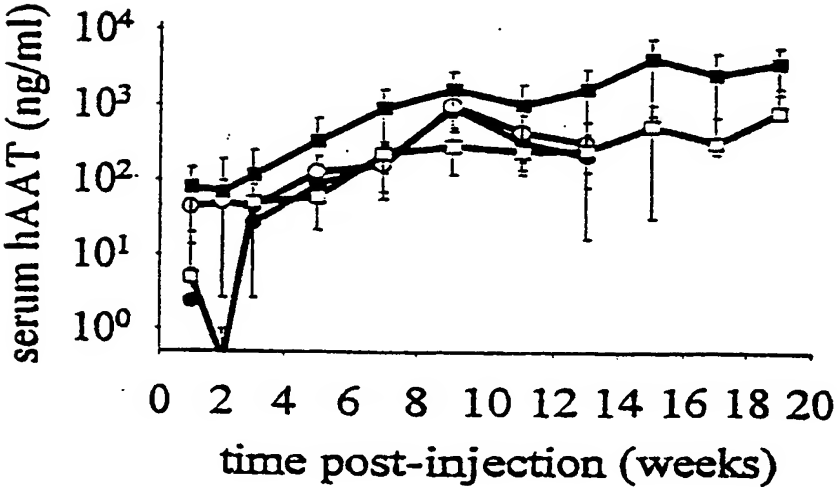


FIGURE 5B

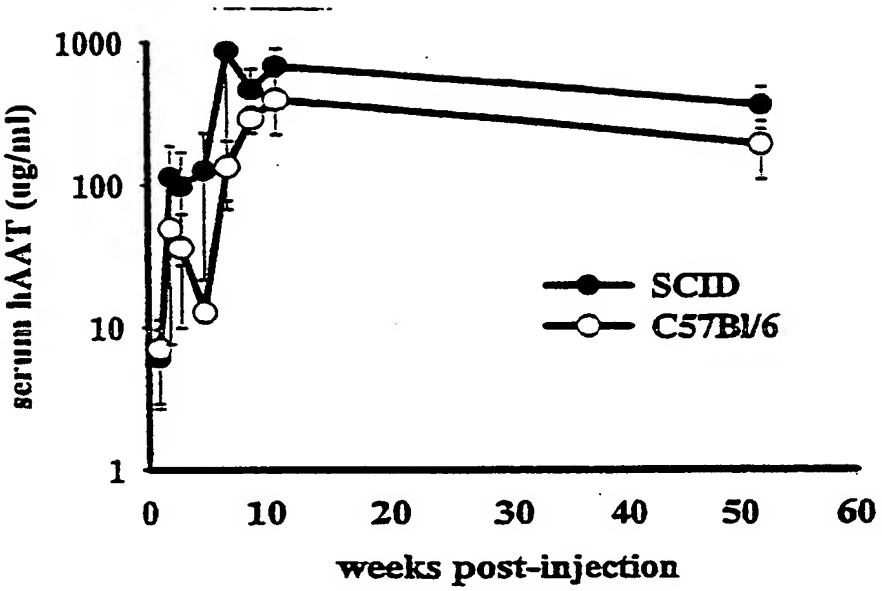

1000
100
10
1
serum hAAT (ug/ml)
0
10
20
30
40
50
60
weeks post-injection
SCID
C57Bl/6

FIGURE 5C

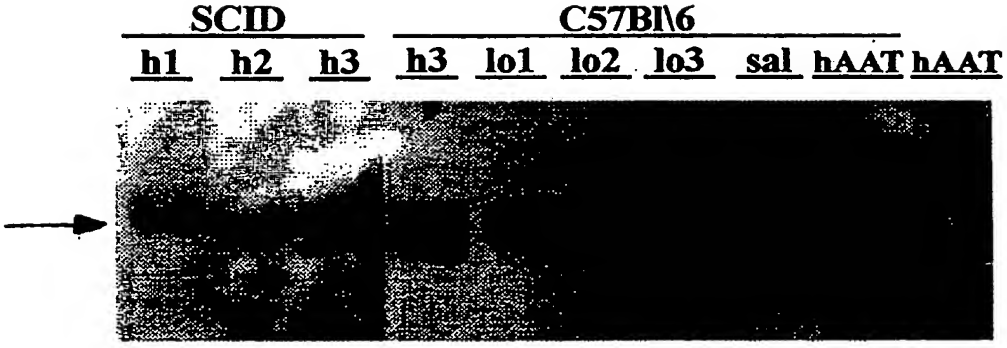


FIGURE 6

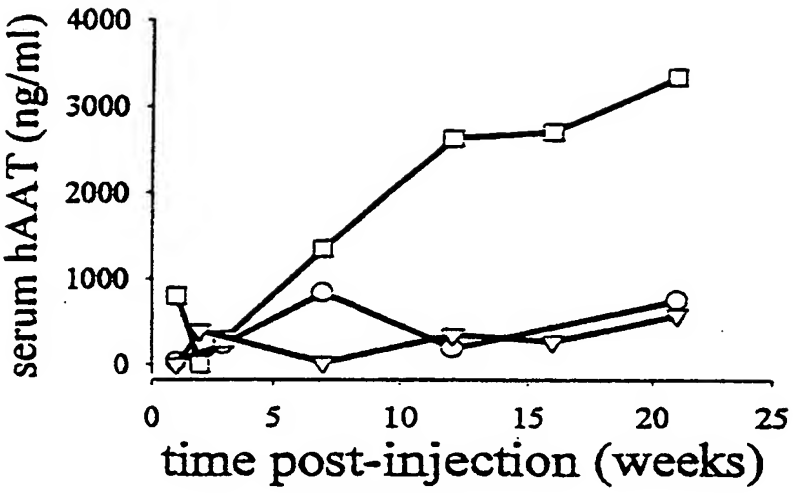


FIGURE 7A

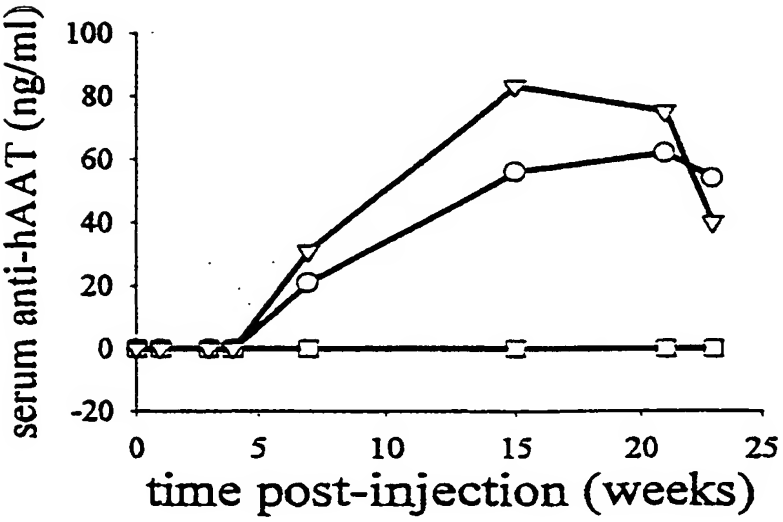


FIGURE 7B

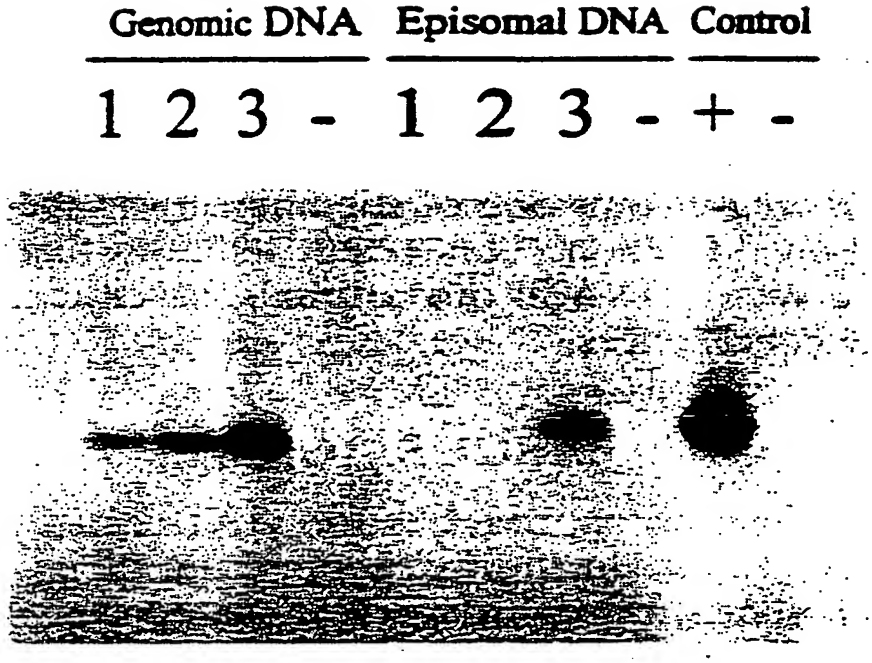


FIGURE 8

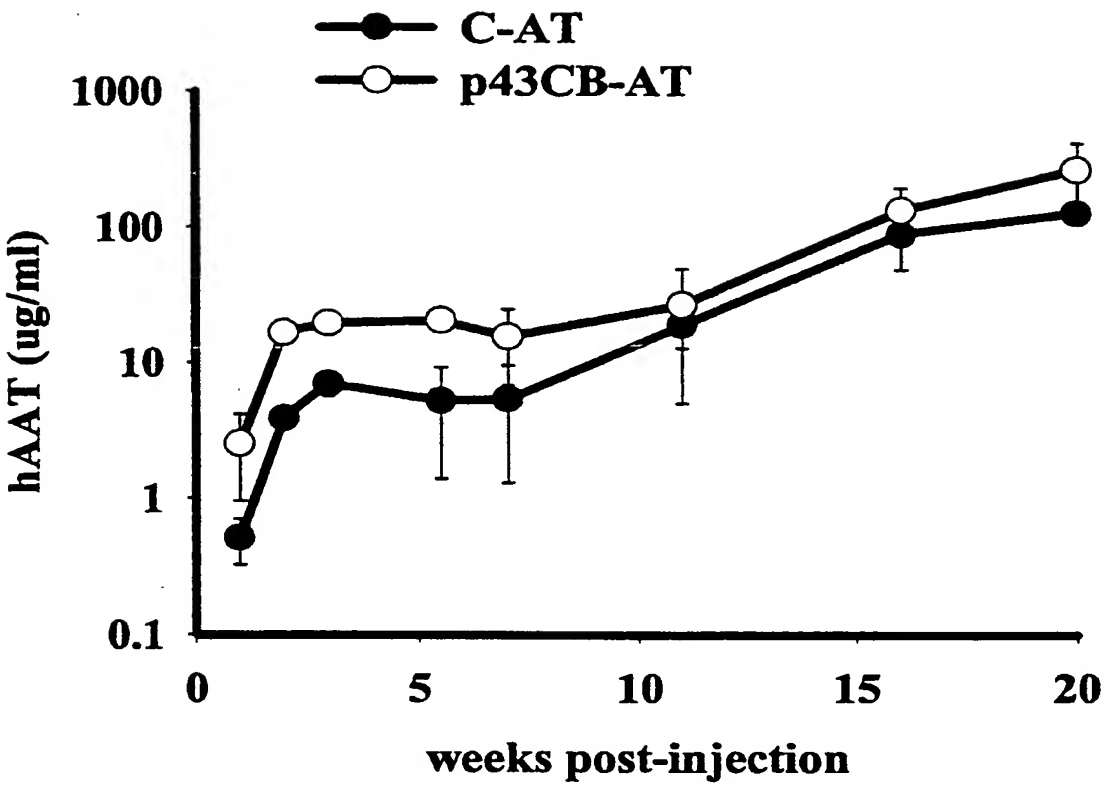

C-AT
p43CB-AT
1000
100
10
1
0.1
hAAT (ug/ml)
0
5
10
15
20
weeks post-injection

FIGURE 9

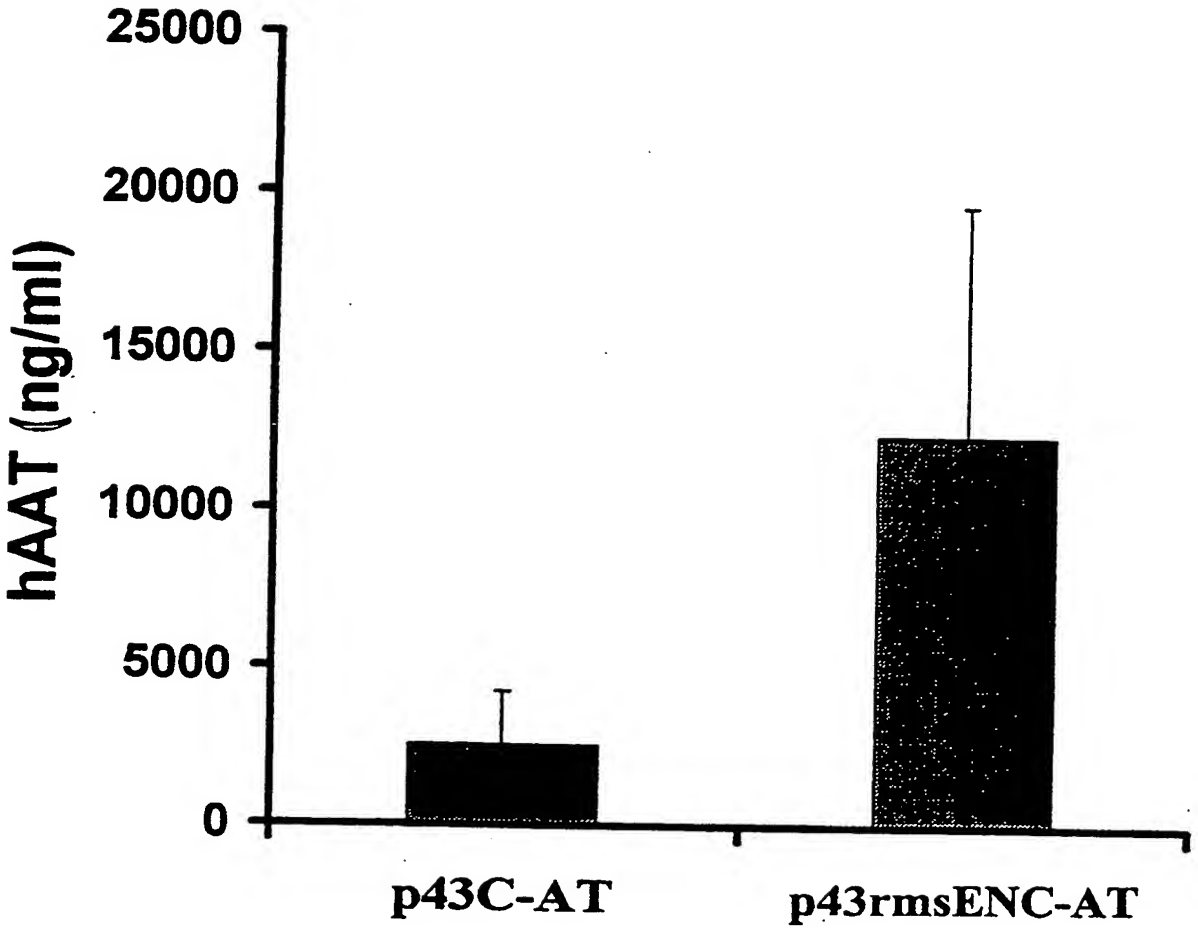

25000
20000
15000
10000
5000
0
hAAT (ng/ml)
p43C-AT
p43rmsENC-AT

FIGURE 10

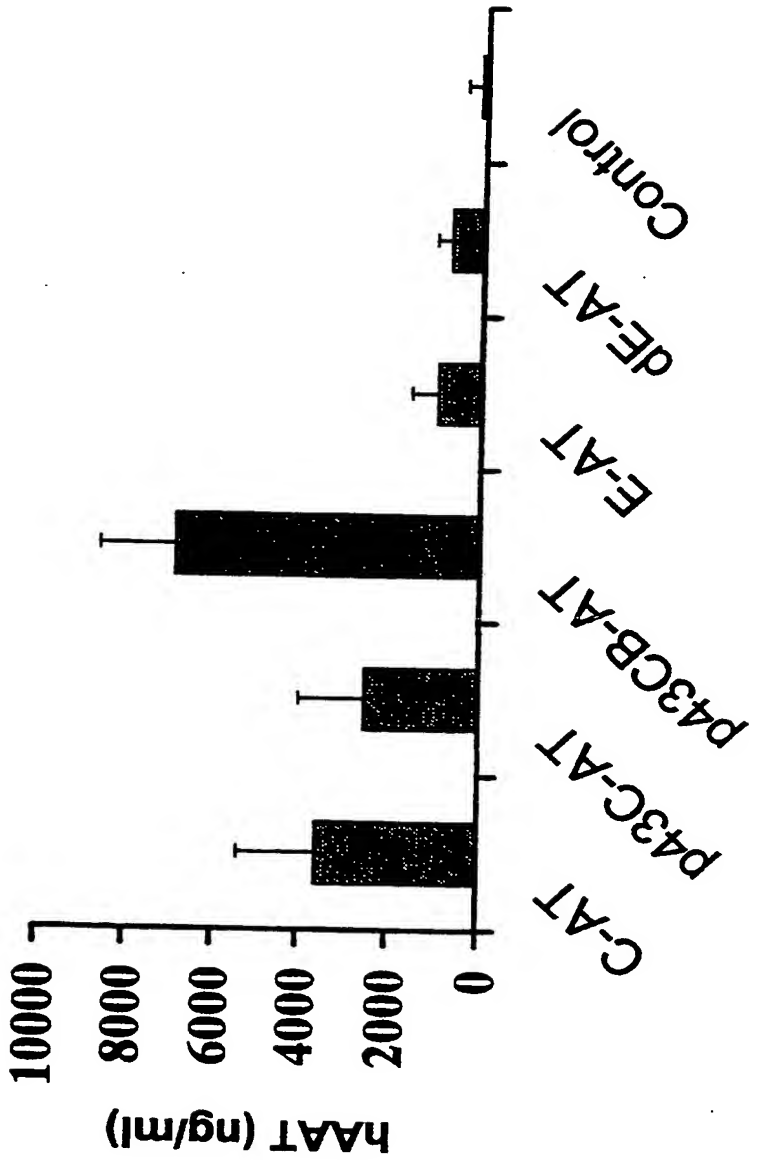

hAAT (ng/ml)
10000
8000
6000
4000
2000
0
C-AT
p43C-AT
p43CB-AT
E-AT
dE-AT
Control

FIGURE 11

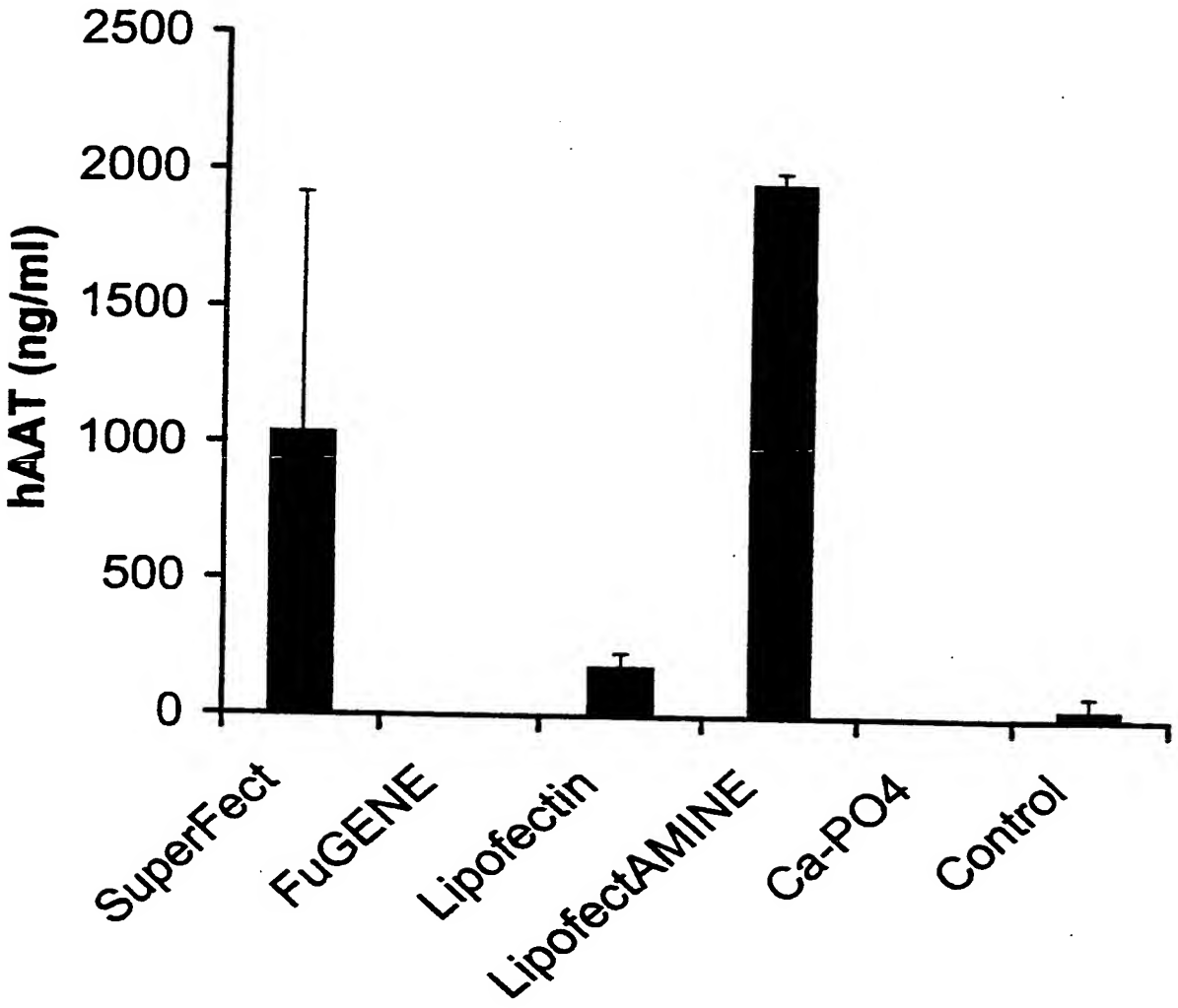

2500
2000
1500
1000
500
0
hAAT (ng/ml)
SuperFect
FuGENE
Lipofectin
LipofectAMINE
Ca-PO4
Control

FIGURE 12

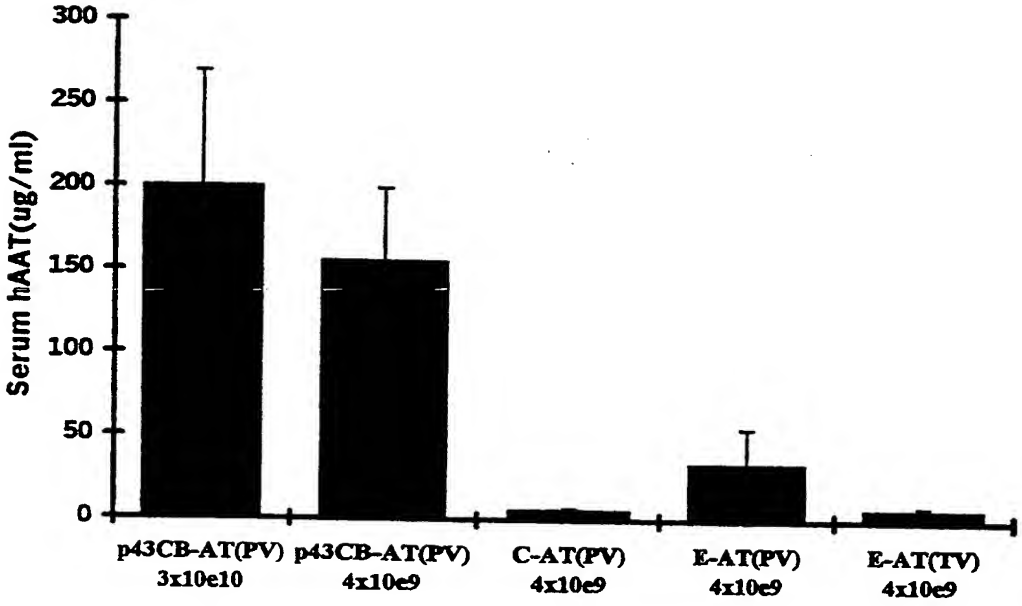

Serum hAAT(ug/ml)
300
250
200
150
100
50
0
p43CB-AT(PV) 3x10e10
p43CB-AT(PV) 4x10e9
C-AT(PV) 4x10e9
E-AT(PV) 4x10e9
E-AT(TV) 4x10e9

FIGURE 13

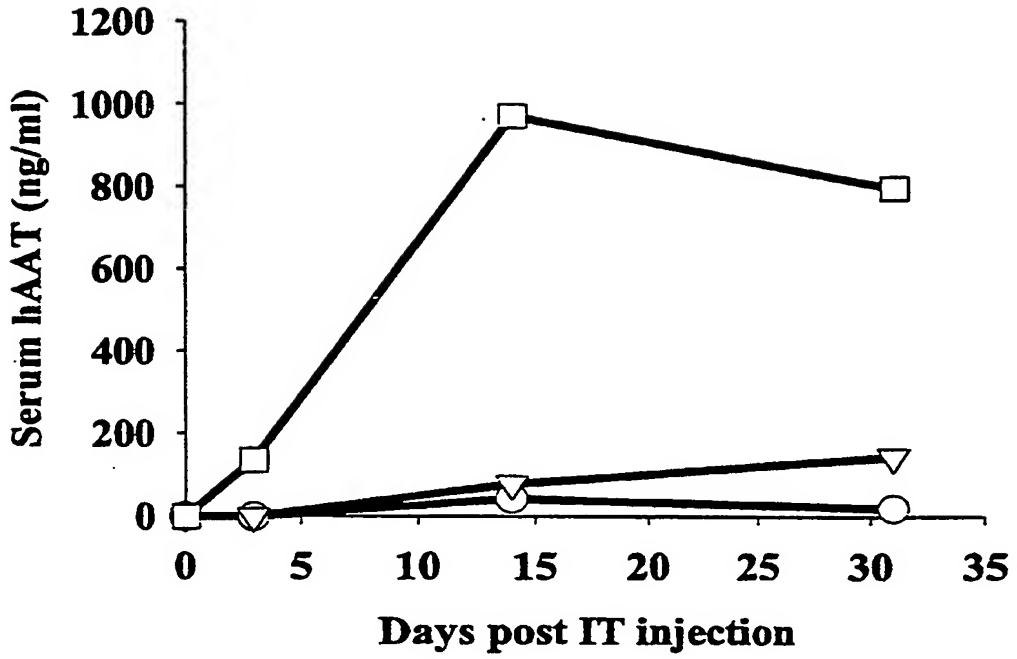

1200
1000
800
600
400
200
0
Serum hAAT (ng/ml)
0
5
10
15
20
25
30
35
Days post IT injection

FIGURE 14

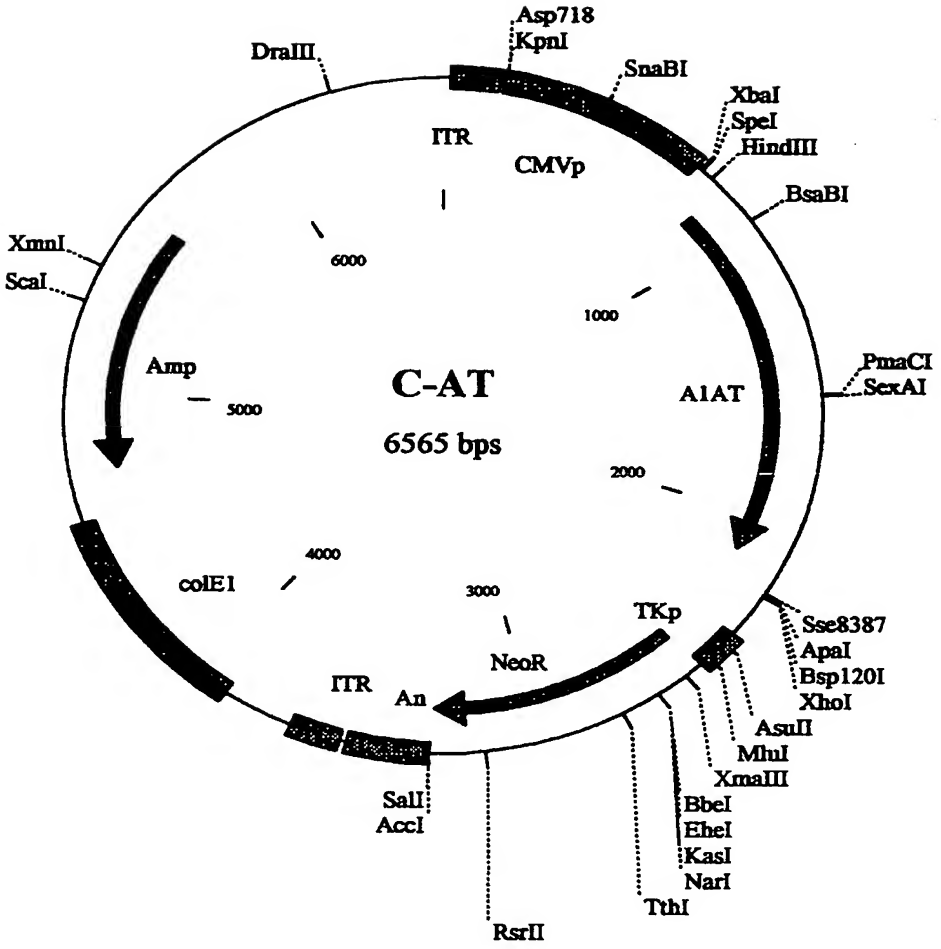

Asp718
KpnI
SnaBI
XbaI
SpeI
HindIII
BsaBI
PmaCI
SexAI
Sse8387
ApaI
Bsp120I
XhoI
AsuII
MluI
XmaIII
BbeI
EheI
KasI
NarI
TthI
RsrII
SalI
AccI
ITR
An
NeoR
TKp
colE1
Amp
XmnI
ScaI
DraIII
ITR
CMVp
A1AT
C-AT
6565 bps
1000
2000
3000
4000
5000
6000

FIGURE 15

Molecule Name: C-AT 6565 bps DNA Circular
Sequence Printed: 1-6565 (Full) Date Printed 16 Apr 1999
Description: Ligation of pTR and aat

```
   1  gggggggggg gggggggttt ggccactccc tctctgcgcg ctcgctcgct
  51  cactgaggcc gggcgaccaa aggtcgcccg acgcccgggc tttgcccggg
 101  cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc
 151  actaggggtt cctagatctg aattcggtac ccgttacata acttacggta
 201  aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat
 251  aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc
 301  aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg
 351  tatcatatgc caagtacgcc ccctattgac gtcaatgacg gtaaatggcc
 401  cgcctggcat tatgcccagt acatgacctt atgggacttt cctacttggc
 451  agtacatcta cgtattagtc atcgctatta ccatggtgat gcggttttgg
 501  cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag
 551  tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac
 601  gggactttcc aaaatgtcgt aacaactccg ccccattgac gcaaatgggc
 651  ggtaggcgtg tacggtggga ggtctatata agcagagctc gtttagtgaa
 701  ccgtcagatc gcctggagac gccatccacg ctgttttgac ctccatagaa
 751  gacaccggga ccgatccagc ctccggactc tagaactagt ggatcccccg
 801  ggctgcagga attcgatatc aagcttgggg attttcaggc accaccactg
 851  acctgggaca gtgaatcgac aatgccgtct tctgtctcgt ggggcatcct
 901  cctgctggca ggcctgtgct gcctggtccc tgtctccctg gctgaggatc
 951  cccagggaga tgctgcccag aagacagata catcccacca tgatcaggat
1001  cacccaacct tcaacaagat caccccccaac ctggctgagt tcgccttcag
1051  cctataccgc cagctggcac accagtccaa cagcaccaat atcttcttct
1101  ccccagtgag catcgctaca gcctttgcaa tgctctccct ggggaccaag
1151  gctgacactc acgatgaaat cctggagggc ctgaatttca acctcacgga
1201  gattccggag gctcagatcc atgaaggctt ccaggaactc ctccgtaccc
1251  tcaaccagcc agacagccag ctccagctga ccaccggcaa tggcctgttc
1301  ctcagcgagg gcctgaagct agtggataag tttttggagg atgttaaaaa
1351  gttgtaccac tcagaagcct tcactgtcaa cttcggggac accgaagagg
1401  ccaagaaaca gatcaacgat tacgtggaga agggtactca agggaaaatt
1451  gtggatttgg tcaaggagct tgacagagac acagtttttg ctctggtgaa
1501  ttacatcttc tttaaaggca aatgggagag acccttttgaa gtcaaggaca
1551  ccgaggaaga ggacttccac gtggaccagg tgaccaccgt gaaggtgcct
1601  atgatgaagc gtttaggcat gtttaacatc cagcactgta agaagctgtc
1651  cagctgggtg ctgctgatga aatacctggg caatgccacc gccatcttct
1701  tcctgcctga tgaggggaaa ctacagcacc tggaaaatga actcacccac
1751  gatatcatca ccaagttcct ggaaaatgaa gacagaaggt ctgccagctt
1801  acatttaccc aaactgtcca ttactggaac ctatgatctg aagagcgtcc
1851  tgggtcaact gggcatcact aaggtcttca gcaatggggc tgacctctcc
1901  ggggtcacag aggaggcacc cctgaagctc tccaaggccg tgcataaggc
1951  tgtgctgacc atcgacgaga aagggactga agctgctggg gccatgtttt
2001  tagaggccat acccatgtct atccccccg aggtcaagtt caacaaaccc
2051  tttgtcttct taatgattga acaaaatacc aagtctcccc tcttcatggg
2101  aaaagtggtg aatcccaccc aaaaataact gcctctcgct cctcaacccc
2151  tcccctccat ccctggcccc ctccctggat gacattaaag aagggttgag
2201  ctggtaaccc ccccccccc tgcaggggcc ctcgagcagt gtggttttgc
2251  aagaggaagc aaaaagcctc tccacccagg cctggaatgt ttccacccaa
2301  gtcgaaggca gtgtggtttt gcaagaggaa gcaaaaagcc tctccaccca
2351  ggcctggaat gtttccaccc aatgtcgagc aaccccgccc agcgtcttgt
2401  cattggcgaa ttcgaacacg cagatgcagt cggggcggcg cggtcccagg
2451  tccacttcgc atattaaggt gacgcgtgtg gcctcgaaca ccgagcgacc
2501  ctgcagccaa tatgggatcg gccattgaac aagatggatt gcacgcaggt
2551  tctccggccg cttgggtgga gaggctattc ggctatgact gggcacaaca
2601  gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca gcgcaggggc
2651  gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg
2701  caggacgagg cagcgcggct atcgtggctg gccacgacgg gcgttccttg
```

FIGURE 15A

```
2751  cgcagctgtg ctcgacgttg tcactgaagc gggaagggac tggctgctat
2801  tgggcgaagt gccggggcag gatctcctgt catctcacct tgctcctgcc
2851  gagaaagtat ccatcatggc tgatgcaatg cggcggctgc atacgcttga
2901  tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcgag
2951  cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa
3001  gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg
3051  catgcccgac ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc
3101  cgaatatcat ggtggaaaat ggccgctttt ctggattcat cgactgtggc
3151  cggctgggtg tggcggaccg ctatcaggac atagcgttgg ctacccgtga
3201  tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctcgtgcttt
3251  acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt
3301  gacgagttct tctgagggga tccgtcgact agagctcgct gatcagcctc
3351  gactgtgcct tctagttgcc agccatctgt tgtttgcccc tcccccgtgc
3401  cttccttgac cctggaaggt gccactccca ctgtcctttc ctaataaaat
3451  gaggaaattg catcgcattg tctgagtagg tgtcattcta ttctgggggg
3501  tggggtgggg caggacagca agggggagga ttgggaagac aatagcaggc
3551  atgctgggga gagatctagg aacccctagt gatggagttg gccactccct
3601  ctctgcgcgc tcgctcgctc actgaggccg cccgggcaaa gcccgggcgt
3651  cgggcgacct ttggtcgccc ggcctcagtg agcgagcgag cgcgcagaga
3701  gggagtggcc aaccccccc cccccccccc tgcagccctg cattaatgaa
3751  tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct
3801  tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt
3851  atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata
3901  acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt
3951  aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga
4001  gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac
4051  tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct
4101  gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg
4151  aagcgtggcg ctttctcaat gctcacgctg taggtatctc agttcggtgt
4201  aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc
4251  gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag
4301  acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag
4351  cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac
4401  ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt
4451  taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg
4501  ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa
4551  aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca
4601  gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa
4651  ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc
4701  taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag
4751  tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct
4801  gactccccgt cgtgtagata actacgatac gggagggctt accatctggc
4851  cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt
4901  atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg
4951  caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga
5001  gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg ccattgctac
5051  aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg
5101  gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa
5151  gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc
5201  agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca
5251  tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca
5301  ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat
5351  acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg
5401  gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga
5451  tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt
5501  tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg
5551  caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc
5601  ctttttcaat attattgaag catttatcag ggttattgtc tcatgagcgg
5651  atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca
5701  catttccccg aaaagtgcca cctgacgtct aagaaaccat tattatcatg
```

FIGURE 15B

```
5751  acattaacct ataaaaatag gcgtatcacg aggccctttc gtctcgcgcg
5801  tttcggtgat gacggtgaaa acctctgaca catgcagctc ccggagacgg
5851  tcacagcttg tctgtaagcg gatgccggga gcagacaagc ccgtcagggc
5901  gcgtcagcgg gtgttggcgg gtgtcggggc tggcttaact atgcggcatc
5951  agagcagatt gtactgagag tgcaccatat gcggtgtgaa ataccgcaca
6001  gatgcgtaag gagaaaatac cgcatcagga aattgtaaac gttaatattt
6051  tgttaaaatt cgcgttaaat ttttgttaaa tcagctcatt ttttaaccaa
6101  taggccgaaa tcggcaaaat cccttataaa tcaaaagaat agaccgagat
6151  agggttgagt gttgttccag tttggaacaa gagtccacta ttaaagaacg
6201  tggactccaa cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca
6251  ctacgtgaac catcacccta atcaagtttt ttggggtcga ggtgccgtaa
6301  agcactaaat cggaacccta aagggagccc ccgatttaga gcttgacggg
6351  gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc gaaaggagcg
6401  ggcgctaggg cgctggcaag tgtagcggtc acgctgcgcg taaccaccac
6451  acccgccgcg cttaatgcgc cgctacaggg cgcgtcgcgc cattcgccat
6501  tcaggctacg caactgttgg gaagggcgat cggtgcgggc ctcttcgcta
6551  ttacgccagg ctgca
```

FIGURE 15C

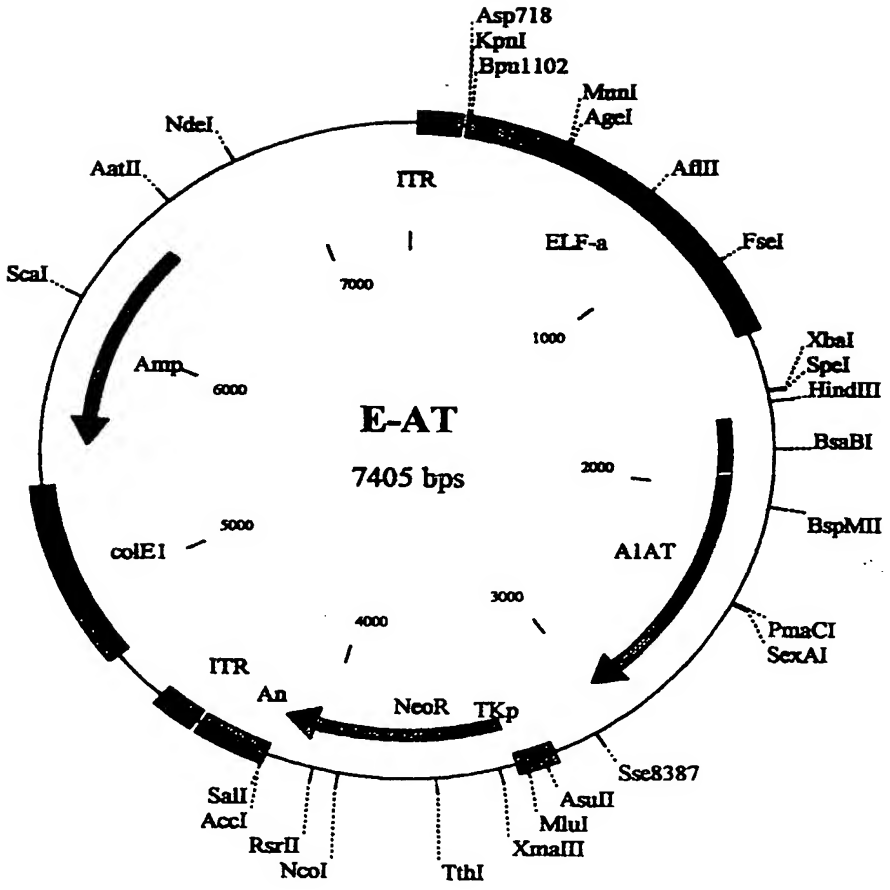

Asp718
KpnI
Bpu1102
MunI
AgeI
AflII
FseI
XbaI
SpeI
HindIII
BsaBI
BspMII
PmaCI
SexAI
Sse8387
AsuII
MluI
XmaIII
TthI
NcoI
RsrII
AccI
SalI
ITR
An
NeoR
TKp
ITR
colE1
ScaI
AatII
NdeI
Amp
E-AT
7405 bps
A1AT
ELF-a
1000
2000
3000
4000
5000
6000
7000

FIGURE 16

Molecule Name: E-AT 7405 bps DNA Circular
Sequence Printed: 1-7405 (Full) Date Printed 16 Apr 1999
Description: Ligation of AAT and elf

```
   1  gggggggggg gggggggttt ggccactccc tctctgcgcg ctcgctcgct
  51  cactgaggcc gggcgaccaa aggtcgcccg acgcccgggc tttgcccggg
 101  cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc
 151  actaggggtt cctagatctg aattcggtac cttggagcta agccagcaat
 201  ggtagaggga agattctgca cgtcccttcc aggcggcctc cccgtcacca
 251  ccccccccaa cccgccccga ccggagctga gagtaattca tacaaaagga
 301  ctcgcccctg ccttggggaa tcccagggac cgtcgttaaa ctcccactaa
 351  cgtagaaccc agagatcgct gcgttcccgc cccctcaccc gcccgctctc
 401  gtcatcactg aggtggagaa gagcatgcgt gaggctccgg tgcccgtcag
 451  tgggcagagc gcacatcgcc cacagtcccc gagaagttgg ggggaggggt
 501  cggcaattga accggtgcct agagaaggtg gcgcggggta aactgggaaa
 551  gtgatgtcgt gtactggctc cgccttttte ccgagggtgg gggagaaccg
 601  tatataagtg cagtagtcgc cgtgaacgtt ctttttcgca acgggtttgc
 651  cgccagaaca caggtaagtg ccgtgtgtgg ttcccgcggg cctggcctct
 701  ttacgggtta tggcccttgc gtgccttgaa ttacttccac gcccctggct
 751  gcagtacgtg attcttgatc ccgagcttcg ggttggaagt gggtgggaga
 801  gttcgaggcc ttgcgcttaa ggagcccctt cgcctcgtgc ttgagttgag
 851  gcctggcctg ggcgctgggg ccgccgcgtg cgaatctggt ggcaccttcg
 901  cgcctgtctc gctgctttcg ataagtctct agccatttaa aatttttgat
 951  gacctgctgc gacgcttttt ttctggcaag atagtcttgt aaatgcgggc
1001  caagatctgc acactggtat ttcggttttt ggggccgcgg gcggcgacgg
1051  ggcccgtgcg tcccagcgca catgttcggc gaggcggggc ctgcgagcgc
1101  ggccaccgag aatcggacgg gggtagtctc aagctggccg gcctgctctg
1151  gtgcctggcc tcgcgccgcc gtgtatcgcc ccgccctggg cggcaaggct
1201  ggcccggtcg gcaccagttg cgtgagcgga aagatggccg cttcccggcc
1251  ctgctgcagg gagctcaaaa tggaggacgc ggcgctcggg agagcgggcg
1301  ggtgagtcac ccacacaaag gaaaagggcc tttccgtcct cagccgtcgc
1351  ttcatgtgac tccacggagt accgggcgcc gtccaggcac ctcgattagt
1401  tctcgagctt ttggagtacg tcgtctttag gttgggggga ggggttttat
1451  gcgatggagt ttccccacac tgagtgggtg gagactgaag ttaggccagc
1501  ttggcacttg atgtaattct ccttggaatt tgccctttt gagtttggat
1551  cttggttcat tctcaagcct cagacagtgg ttcaaagttt ttttcttcca
1601  tttcaggtgt cgtgaaaatc tagaactagt ggatcccccg ggctgcagga
1651  attcgatatc aagcttgggg attttcaggc accaccactg acctgggaca
1701  gtgaatcgac aatgccgtct tctgtctcgt ggggcatcct cctgctggca
1751  ggcctgtgct gcctggtccc tgtctccctg gctgaggatc cccagggaga
1801  tgctgcccag aagacagata catcccacca tgatcaggat cacccaacct
1851  tcaacaagat caccccaac ctggctgagt tcgccttcag cctataccgc
1901  cagctggcac accagtccaa cagcaccaat atcttcttct ccccagtgag
1951  catcgctaca gcctttgcaa tgctctccct ggggaccaag gctgacactc
2001  acgatgaaat cctggagggc ctgaatttca acctcacgga gattccggag
2051  gctcagatcc atgaaggctt ccaggaactc ctccgtaccc tcaaccagcc
2101  agacagccag ctccagctga ccaccggcaa tggcctgttc ctcagcgagg
2151  gcctgaagct agtggataag tttttggagg atgttaaaaa gttgtaccac
2201  tcagaagcct tcactgtcaa cttcggggac accgaagagg ccaagaaaca
2251  gatcaacgat tacgtggaga agggtactca agggaaaatt gtggatttgg
2301  tcaaggagct tgacagagac acagtttttg ctctggtgaa ttacatcttc
2351  tttaaaggca aatgggagag acccttttgaa gtcaaggaca ccgaggaaga
2401  ggacttccac gtggaccagg tgaccaccgt gaaggtgcct atgatgaagc
2451  gtttaggcat gtttaacatc cagcactgta agaagctgtc cagctgggtg
2501  ctgctgatga aatacctggg caatgccacc gccatcttct tcctgcctga
2551  tgaggggaaa ctacagcacc tggaaaatga actcacccac gatatcatca
2601  ccaagttcct ggaaaatgaa gacagaaggt ctgccagctt acatttaccc
2651  aaactgtcca ttactggaac ctatgatctg aagagcgtcc tgggtcaact
2701  gggcatcact aaggtcttca gcaatggggc tgacctctcc ggggtcacag
```

FIGURE 16A

```
2751  aggaggcacc ccaagctc tccaaggccg tgcataaggc tgcgctgacc
2801  atcgacgaga aagggactga agctgctggg gccatgtttt tagaggccat
2851  acccatgtct atccccccg aggtcaagtt caacaaaccc tttgtcttct
2901  taatgattga acaaaatacc aagtctcccc tcttcatggg aaaagtggtg
2951  aatcccaccc aaaaataact gcctctcgct cctcaacccc tcccctccat
3001  ccctggcccc ctccctggat gacattaaag aagggttgag ctggtaaccc
3051  ccccccccc tgcaggggcc ctcgagcagt gtggttttgc aagaggaagc
3101  aaaaagcctc tccacccagg cctggaatgt ttccacccaa gtcgaaggca
3151  gtgtggtttt gcaagaggaa gcaaaaagcc tctccaccca ggcctggaat
3201  gtttccaccc aatgtcgagc aaccccgccc agcgtcttgt cattggcgaa
3251  ttcgaacacg cagatgcagt cggggcggcg cggtcccagg tccacttcgc
3301  atattaaggt gacgcgtgtg gcctcgaaca ccgagcgacc ctgcagccaa
3351  tatgggatcg gccattgaac aagatggatt gcacgcaggt tctccggccg
3401  cttgggtgga gaggctattc ggctatgact gggcacaaca gacaatcggc
3451  tgctctgatg ccgccgtgtt ccggctgtca gcgcaggggc gcccggttct
3501  ttttgtcaag accgacctgt ccggtgccct gaatgaactg caggacgagg
3551  cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg
3601  ctcgacgttg tcactgaagc gggaagggac tggctgctat tgggcgaagt
3651  gccggggcag gatctcctgt catctcacct tgctcctgcc gagaaagtat
3701  ccatcatggc tgatgcaatg cggcggctgc atacgcttga tccggctacc
3751  tgcccattcg accaccaagc gaaacatcgc atcgagcgag cacgtactcg
3801  gatggaagcc ggtcttgtcg atcaggatga tctggacgaa gagcatcagg
3851  ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg catgcccgac
3901  ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat
3951  ggtggaaaat ggccgctttt ctggattcat cgactgtggc cggctgggtg
4001  tggcggaccg ctatcaggac atagcgttgg ctacccgtga tattgctgaa
4051  gagcttggcg gcgaatgggc tgaccgcttc ctcgtgcttt acggtatcgc
4101  cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagttct
4151  tctgagggga tccgtcgact agagctcgct gatcagcctc gactgtgcct
4201  tctagttgcc agccatctgt tgtttgcccc tcccccgtgc cttccttgac
4251  cctggaaggt gccactccca ctgtcctttc ctaataaaat gaggaaattg
4301  catcgcattg tctgagtagg tgtcattcta ttctgggggg tggggtgggg
4351  caggacagca agggggagga ttgggaagac aatagcaggc atgctgggga
4401  gagatctagg aacccctagt gatggagttg gccactccct ctctgcgcgc
4451  tcgctcgctc actgaggccg cccgggcaaa gcccgggcgt cgggcgacct
4501  ttggtcgccc ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc
4551  aacccccccc ccccccccc tgcagccctg cattaatgaa tcggccaacg
4601  cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca
4651  ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac
4701  tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa
4751  gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg
4801  cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa
4851  aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata
4901  ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc
4951  tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg
5001  ctttctcaat gctcacgctg taggtatctc agttcggtgt aggtcgttcg
5051  ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg
5101  ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta
5151  tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt
5201  aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta
5251  gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga
5301  aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg
5351  tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc
5401  aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa
5451  aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac
5501  ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat
5551  atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct
5601  atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt
5651  cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg
5701  caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata
```

FIGURE 16B

```
5751  aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc
5801  cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt
5851  cgccagttaa tagtttgcgc aacgttgttg ccattgctac aggcatcgtg
5901  gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg gttcccaacg
5951  atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct
6001  ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca
6051  ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt
6101  aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat
6151  agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat
6201  accgcgccac atagcagaac tttaaaagtg ctcatcattg gaaaacgttc
6251  ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga
6301  tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc
6351  agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg
6401  aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat
6451  attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt
6501  gaatgtattt agaaaaataa acaaataggg gttccgcgca catttccccg
6551  aaaagtgcca cctgacgtct aagaaaccat tattatcatg acattaacct
6601  ataaaaatag gcgtatcacg aggccctttc gtctcgcgcg tttcggtgat
6651  gacggtgaaa acctctgaca catgcagctc ccggagacgg tcacagcttg
6701  tctgtaagcg gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg
6751  gtgttggcgg gtgtcggggc tggcttaact atgcggcatc agagcagatt
6801  gtactgagag tgcaccatat gcggtgtgaa ataccgcaca gatgcgtaag
6851  gagaaaatac cgcatcagga aattgtaaac gttaatattt tgttaaaatt
6901  cgcgttaaat ttttgttaaa tcagctcatt ttttaaccaa taggccgaaa
6951  tcggcaaaat cccttataaa tcaaaagaat agaccgagat agggttgagt
7001  gttgttccag tttggaacaa gagtccacta ttaaagaacg tggactccaa
7051  cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca ctacgtgaac
7101  catcacccta atcaagtttt ttggggtcga ggtgccgtaa agcactaaat
7151  cggaacccta aagggagccc ccgatttaga gcttgacggg gaaagccggc
7201  gaacgtggcg agaaaggaag ggaagaaagc gaaaggagcg ggcgctaggg
7251  cgctggcaag tgtagcggtc acgctgcgcg taaccaccac acccgccgcg
7301  cttaatgcgc cgctacaggg cgcgtcgcgc cattcgccat tcaggctacg
7351  caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagg
7401  ctgca
```

FIGURE 16C

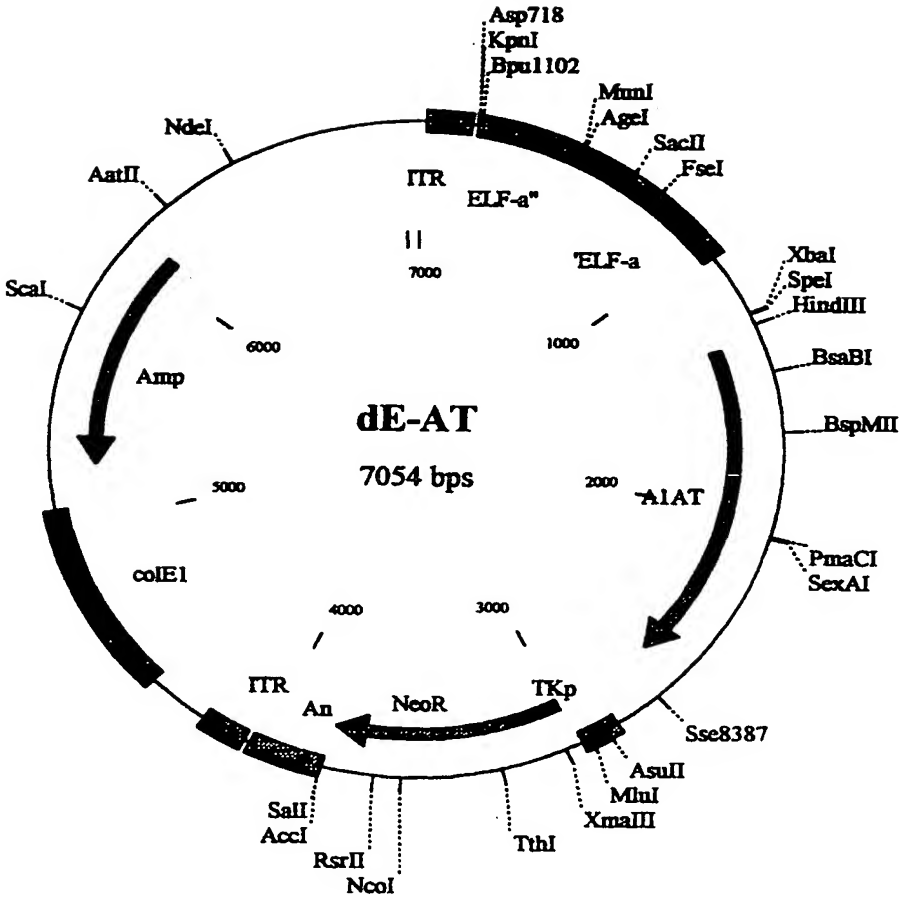

Asp718
KpnI
Bpu1102
MunI
AgeI
SacII
FseI
XbaI
SpeI
HindIII
BsaBI
BspMII
PmaCI
SexAI
Sse8387
AsuII
MluI
XmaIII
TthI
SalI
AccI
RsrII
NcoI
ITR
An
NeoR
TKp
colE1
ScaI
AatII
NdeI
Amp
ITR
ELF-a"
ELF-a
A1AT
dE-AT
7054 bps
1000
2000
3000
4000
5000
6000
7000

FIGURE 17

Molecule Name: dE-AT 7054 bps DNA Circular
Sequence Printed: 1-7054 (Full) Date Printed 16 Apr 1999
Description: Fragment 2 Circularized

```
   1  gggggggggg gggggggttt ggccactccc tctctgcgcg ctcgctcgct
  51  cactgaggcc gggcgaccaa aggtcgcccg acgcccgggc tttgcccggg
 101  cggcctcagt gagcgagcga gcgcgcagag agggagtggc caactccatc
 151  actaggggtt cctagatctg aattcggtac cttggagcta agccagcaat
 201  ggtagaggga agattctgca cgtcccttcc aggcggcctc cccgtcacca
 251  ccccccccaa cccgccccga ccggagctga gagtaattca tacaaaagga
 301  ctcgcccctg ccttggggaa tcccagggac cgtcgttaaa ctcccactaa
 351  cgtagaaccc agagatcgct gcgttcccgc cccctcaccc gcccgctctc
 401  gtcatcactg aggtggagaa gagcatgcgt gaggctccgg tgcccgtcag
 451  tgggcagagc gcacatcgcc cacagtcccc gagaagttgg ggggagggggt
 501  cggcaattga accggtgcct agagaaggtg gcgcggggta aactgggaaa
 551  gtgatgtcgt gtactggctc cgccttttc ccgagggtgg gggagaaccg
 601  tatataagtg cagtagtcgc cgtgaacgtt ctttttcgca acgggtttgc
 651  cgccagaaca caggtaagtg ccgtgtgtgg ttcccgcggg cggcgacggg
 701  gcccgtgcgt cccagcgcac atgttcggcg aggcggggcc tgcgagcgcg
 751  gccaccgaga atcggacggg ggtagtctca agctggccgg cctgctctgg
 801  tgcctggcct cgcgccgccg tgtatcgccc cgccctgggc ggcaaggctg
 851  gcccggtcgg caccagttgc gtgagcggaa agatggccgc ttcccggccc
 901  tgctgcaggg agctcaaaat ggaggacgcg gcgctcggga gagcgggcgg
 951  gtgagtcacc cacacaaagg aaaagggcct ttccgtcctc agccgtcgct
1001  tcatgtgact ccacggagta ccgggcgccg tccaggcacc tcgattagtt
1051  ctcgagcttt tggagtacgt cgtctttagg ttgggggggag gggttttatg
1101  cgatggagtt tccccacact gagtgggtgg agactgaagt taggccagct
1151  tggcacttga tgtaattctc cttggaattt gccctttttg agtttggatc
1201  ttggttcatt ctcaagcctc agacagtggt tcaaagtttt tttcttccat
1251  ttcaggtgtc gtgaaaatct agaactagtg gatcccccgg gctgcaggaa
1301  ttcgatatca agcttgggga ttttcaggca ccaccactga cctgggacag
1351  tgaatcgaca atgccgtctt ctgtctcgtg gggcatcctc ctgctggcag
1401  gcctgtgctg cctggtccct gtctccctgg ctgaggatcc ccagggagat
1451  gctgcccaga agacagatac atcccaccat gatcaggatc acccaacctt
1501  caacaagatc acccccaacc tggctgagtt cgccttcagc ctataccgcc
1551  agctggcaca ccagtccaac agcaccaata tcttcttctc cccagtgagc
1601  atcgctacag cctttgcaat gctctccctg gggaccaagg ctgacactca
1651  cgatgaaatc ctggagggcc tgaatttcaa cctcacggag attccggagg
1701  ctcagatcca tgaaggcttc caggaactcc tccgtaccct caaccagcca
1751  gacagccagc tccagctgac caccggcaat ggcctgttcc tcagcgaggg
1801  cctgaagcta gtggataagt ttttggagga tgttaaaaag ttgtaccact
1851  cagaagcctt cactgtcaac ttcggggaca ccgaagaggc caagaaacag
1901  atcaacgatt acgtggagaa gggtactcaa gggaaaattg tggatttggt
1951  caaggagctt gacagagaca cagtttttgc tctggtgaat tacatcttct
2001  ttaaaggcaa atgggagaga ccctttgaag tcaaggacac cgaggaagag
2051  gacttccacg tggaccaggt gaccaccgtg aaggtgccta tgatgaagcg
2101  tttaggcatg tttaacatcc agcactgtaa gaagctgtcc agctgggtgc
2151  tgctgatgaa atacctgggc aatgccaccg ccatcttctt cctgcctgat
2201  gaggggaaac tacagcacct ggaaaatgaa ctcacccacg atatcatcac
2251  caagttcctg gaaaatgaag acagaaggtc tgccagctta catttaccca
2301  aactgtccat tactggaacc tatgatctga agagcgtcct gggtcaactg
2351  ggcatcacta aggtcttcag caatggggct gacctctccg gggtcacaga
2401  ggaggcaccc ctgaagctct ccaaggccgt gcataaggct gtgctgacca
2451  tcgacgagaa agggactgaa gctgctgggg ccatgttttt agaggccata
2501  cccatgtcta tccccccga ggtcaagttc aacaaaccct ttgtcttctt
2551  aatgattgaa caaaatacca agtctcccct cttcatggga aaagtggtga
2601  atcccaccca aaaataactg cctctcgctc ctcaacccct cccctccatc
2651  cctggccccc tccctggatg acattaaaga agggttgagc tggtaacccc
2701  ccccccccct gcaggggccc tcgagcagtg tggttttgca agaggaagca
```

FIGURE 17A

```
2751  aaaagcctct ccacccaggc ctggaatgtt tccacccaag tcgaaggcag
2801  tgtggttttg caagaggaag caaaaagcct ctccacccag gcctggaatg
2851  tttccaccca atgtcgagca accccgccca gcgtcttgtc attggcgaat
2901  tcgaacacgc agatgcagtc ggggcggcgc ggtcccaggt ccacttcgca
2951  tattaaggtg acgcgtgtgg cctcgaacac cgagcgaccc tgcagccaat
3001  atgggatcgg ccattgaaca agatggattg cacgcaggtt ctccggccgc
3051  ttgggtggag aggctattcg gctatgactg ggcacaacag acaatcggct
3101  gctctgatgc cgccgtgttc cggctgtcag cgcaggggcg cccggttctt
3151  tttgtcaaga ccgacctgtc cggtgccctg aatgaactgc aggacgaggc
3201  agcgcggcta tcgtggctgg ccacgacggg cgttccttgc gcagctgtgc
3251  tcgacgttgt cactgaagcg ggaagggact ggctgctatt gggcgaagtg
3301  ccggggcagg atctcctgtc atctcacctt gctcctgccg agaaagtatc
3351  catcatggct gatgcaatgc ggcggctgca tacgcttgat ccggctacct
3401  gcccattcga ccaccaagcg aaacatcgca tcgagcgagc acgtactcgg
3451  atggaagccg gtcttgtcga tcaggatgat ctggacgaag agcatcaggg
3501  gctcgcgcca gccgaactgt tcgccaggct caaggcgcgc atgcccgacg
3551  gcgaggatct cgtcgtgacc catggcgatg cctgcttgcc gaatatcatg
3601  gtggaaaatg gccgcttttc tggattcatc gactgtggcc ggctgggtgt
3651  ggcggaccgc tatcaggaca tagcgttggc tacccgtgat attgctgaag
3701  agcttggcgg cgaatgggct gaccgcttcc tcgtgcttta cggtatcgcc
3751  gctcccgatt cgcagcgcat cgccttctat cgccttcttg acgagttctt
3801  ctgaggggat ccgtcgacta gagctcgctg atcagcctcg actgtgcctt
3851  ctagttgcca gccatctgtt gtttgccccc cccccgtgcc ttccttgacc
3901  ctggaaggtg ccactcccac tgtcctttcc taataaaatg aggaaattgc
3951  atcgcattgt ctgagtaggt gtcattctat tctgggggggt ggggtggggc
4001  aggacagcaa gggggaggat tgggaagaca atagcaggca tgctggggag
4051  agatctagga acccctagtg atggagttgg ccactccctc tctgcgcgct
4101  cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc gggcgacctt
4151  tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca
4201  acccccccccc ccccccccct gcagccctgc attaatgaat cggccaacgc
4251  gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac
4301  tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact
4351  caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag
4401  aacatgtgag caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc
4451  gttgctggcg tttttccata ggctccgccc ccctgacgag catcacaaaa
4501  atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac
4551  caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct
4601  gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc
4651  tttctcaatg ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc
4701  tccaagctgg gctgtgtgca cgaacccccc gttcagcccg accgctgcgc
4751  cttatccggt aactatcgtc ttgagtccaa cccggtaaga cacgacttat
4801  cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta
4851  ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag
4901  aaggacagta tttggtatct gcgctctgct gaagccagtt accttcggaa
4951  aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt
5001  ggtttttttg tttgcaagca gcagattacg cgcagaaaaa aaggatctca
5051  agaagatcct ttgatctttt ctacggggtc tgacgctcag tggaacgaaa
5101  actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc
5151  tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata
5201  tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta
5251  tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc
5301  gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgctgc
5351  aatgataccg cgagacccac gctcaccggc tccagattta tcagcaataa
5401  accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc
5451  gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc
5501  gccagttaat agtttgcgca acgttgttgc cattgctaca ggcatcgtgg
5551  tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga
5601  tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc
5651  cttcggtcct ccgatcgttg tcagaagtaa gttggccgca gtgttatcac
5701  tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta
```

FIGURE 17B

```
5751  agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata
5801  gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata
5851  ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct
5901  tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat
5951  gtaacccact cgtgcaccca actgatcttc agcatctttt actttcacca
6001  gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga
6051  ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata
6101  ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg
6151  aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga
6201  aaagtgccac ctgacgtcta agaaaccatt attatcatga cattaaccta
6251  taaaaatagg cgtatcacga ggccctttcg tctcgcgcgt ttcggtgatg
6301  acggtgaaaa cctctgacac atgcagctcc cggagacggt cacagcttgt
6351  ctgtaagcgg atgccgggag cagacaagcc cgtcagggcg cgtcagcggg
6401  tgttggcggg tgtcggggct ggcttaacta tgcggcatca gagcagattg
6451  tactgagagt gcaccatatg cggtgtgaaa taccgcacag atgcgtaagg
6501  agaaaatacc gcatcaggaa attgtaaacg ttaatatttt gttaaaattc
6551  gcgttaaatt tttgttaaat cagctcattt tttaaccaat aggccgaaat
6601  cggcaaaatc ccttataaat caaaagaata gaccgagata gggttgagtg
6651  ttgttccagt ttggaacaag agtccactat taaagaacgt ggactccaac
6701  gtcaaagggc gaaaaaccgt ctatcagggc gatggcccac tacgtgaacc
6751  atcaccctaa tcaagttttt tggggtcgag gtgccgtaaa gcactaaatc
6801  ggaaccctaa agggagcccc cgatttagag cttgacgggg aaagccggcg
6851  aacgtggcga gaaaggaagg gaagaaagcg aaaggagcgg gcgctagggc
6901  gctggcaagt gtagcggtca cgctgcgcgt aaccaccaca cccgccgcgc
6951  ttaatgcgcc gctacagggc gcgtcgcgcc attcgccatt caggctacgc
7001  aactgttggg aagggcgatc ggtgcgggcc tcttcgctat tacgccaggc
7051  tgca
```

FIGURE 17C

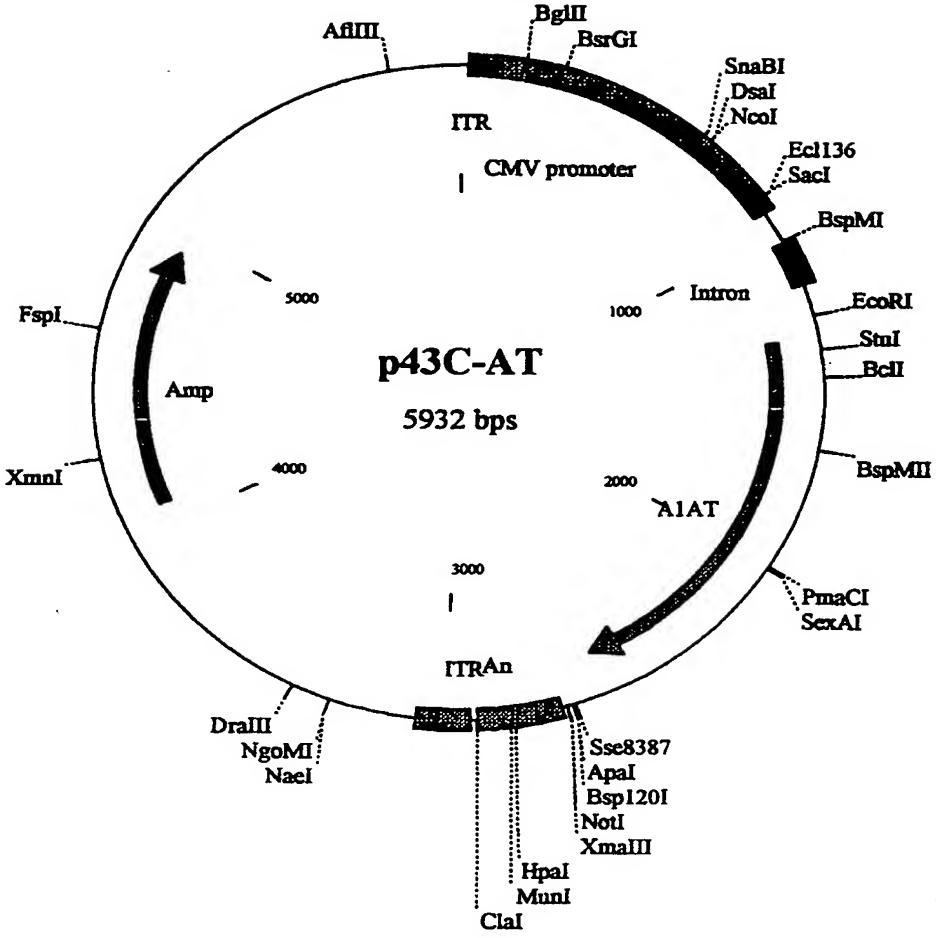

AflIII
BglII
BsrGI
SnaBI
DsaI
NcoI
Ecl136
SacI
BspMI
EcoRI
StuI
BclI
BspMII
PmaCI
SexAI
PflMI
SspI
DraIII
NgoMI
NaeI
Sse8387
ApaI
Bsp120I
NotI
XmaIII
HpaI
MunI
ClaI
XmnI
FspI
ITR
CMV promoter
Intron
A1AT
ITRAn
Amp
p43C-AT
5932 bps
1000
2000
3000
4000
5000

FIGURE 18

Molecule Name: p43C-AT 5932 bps DNA Circular
Sequence Printed: 1-5932 (Full) Date Printed 16 Apr 1999
Description: Ligation of TR and aat

```
   1  gggggggggg gggggggttg gccactccct ctctgcgcgc tcgctcgctc
  51  actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc
 101  ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca
 151  ctaggggttc ctagatcttc aatattggcc attagccata ttattcattg
 201  gttatatagc ataaatcaat attggctatt ggccattgca tacgttgtat
 251  ctatatcata atatgtacat ttatattggc tcatgtccaa tatgaccgcc
 301  atgttggcat tgattattga ctagttatta atagtaatca attacggggt
 351  cattagttca tagcccatat atggagttcc gcgttacata acttacggta
 401  aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat
 451  aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc
 501  aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg
 551  tatcatatgc caagtccgcc ccctattgac gtcaatgacg gtaaatggcc
 601  cgcctggcat tatgcccagt acatgacctt acgggacttt cctacttggc
 651  agtacatcta cgtattagtc atcgctatta ccatggtgat gcggttttgg
 701  cagtacacca atgggcgtgg atagcggttt gactcacggg gatttccaag
 751  tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac
 801  gggactttcc aaaatgtcgt aataaccccg ccccgttgac gcaaatgggc
 851  ggtaggcgtg tacggtggga ggtctatata agcagagctc gtttagtgaa
 901  ccgtcagatc actagaagct ttattgcggt agtttatcac agttaaattg
 951  ctaacgcagt cagtgcttct gacacaacag tctcgaactt aagctgcaga
1001  agttggtcgt gaggcactgg gcaggtaagt atcaaggtta caagacaggt
1051  ttaaggagac caatagaaac tgggcttgtc gagacagaga agactcttgc
1101  gtttctgata ggcacctatt ggtcttactg acatccactt tgcctttctc
1151  tccacaggtg tccactccca gttcaattac agctcttaag gctagagtac
1201  ttaatacgac tcactatagg ctagaactag tggatccccc gggctgcagg
1251  aattcgatat caagcttggg gattttcagg caccaccact gacctgggac
1301  agtgaatcga caatgccgtc ttctgtctcg tggggcatcc tcctgctggc
1351  aggcctgtgc tgcctggtcc ctgtctccct ggctgaggat ccccagggag
1401  atgctgccca gaagacagat acatcccacc atgatcagga tcacccaacc
1451  ttcaacaaga tcaccccaa cctggctgag ttcgccttca gcctataccg
1501  ccagctggca caccagtcca acagcaccaa tatcttcttc tccccagtga
1551  gcatcgctac agcctttgca atgctctccc tggggaccaa ggctgacact
1601  cacgatgaaa tcctggaggg cctgaatttc aacctcacgg agattccgga
1651  ggctcagatc catgaaggct tccaggaact cctccgtacc ctcaaccagc
1701  cagacagcca gctccagctg accaccggca atggcctgtt cctcagcgag
1751  ggcctgaagc tagtggataa gtttttggag gatgttaaaa agttgtacca
1801  ctcagaagcc ttcactgtca acttcgggga caccgaagag gccaagaaac
1851  agatcaacga ttacgtggag aagggtactc aagggaaaat tgtggatttg
1901  gtcaaggagc ttgacagaga cacagttttt gctctggtga attacatctt
1951  ctttaaaggc aaatgggaga gacccttga agtcaaggac accgaggaag
2001  aggacttcca cgtggaccag gtgaccaccg tgaaggtgcc tatgatgaag
2051  cgtttaggca tgtttaacat ccagcactgt aagaagctgt ccagctgggt
2101  gctgctgatg aaatacctgg gcaatgccac cgccatcttc ttcctgcctg
2151  atgaggggaa actacagcac ctggaaaatg aactcaccca cgatatcatc
2201  accaagttcc tggaaaatga agacagaagg tctgccagct tacatttacc
2251  caaactgtcc attactggaa cctatgatct gaagagcgtc ctgggtcaac
2301  tgggcatcac taaggtcttc agcaatgggg ctgacctctc cggggtcaca
2351  gaggaggcac ccctgaagct ctccaaggcc gtgcataagg ctgtgctgac
2401  catcgacgag aaagggactg aagctgctgg ggccatgttt ttagaggcca
2451  tacccatgtc tatccccccc gaggtcaagt tcaacaaacc ctttgtcttc
2501  ttaatgattg aacaaaatac caagtctccc ctcttcatgg gaaaagtggt
2551  gaatcccacc caaaaataac tgcctctcgc tcctcaaccc ctccccctcca
2601  tccctggccc cctccctgga tgacattaaa gaagggttga gctggtaacc
2651  ccccccccc ctgcaggggc cctcgacccg ggcggccgct tcgagcagac
2701  atgataagat acattgatga gtttggacaa accacaacta gaatgcagtg
```

FIGURE 18A

```
2751  aaaaaaatgc tttatttgtg aaatttgtga tgctattgct ttatttgtaa
2801  ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt
2851  atgtttcagg ttcaggggga gatgtgggag gttttttaaa gcaagtaaaa
2901  cctctacaaa tgtggtaaaa tcgataagga tctaggaacc cctagtgatg
2951  gagttggcca ctccctctct gcgcgctcgc tcgctcactg aggccgcccg
3001  ggcaaagccc gggcgtcggg cgacctttgg tcgcccggcc tcagtgagcg
3051  agcgagcgcg cagagaggga gtggccaacc ccccccccgc ccccctgca
3101  gcctggcgta atagcgaaga ggcccgcacc gatcgccctt cccaacagtt
3151  gcgtagcctg aatggcgaat ggcgcgacgc gccctgtagc ggcgcattaa
3201  gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac acttgccagc
3251  gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt
3301  cgccggcttt ccccgtcaag ctctaaatcg ggggctccct ttagggttcc
3351  gatttagtgc tttacggcac ctcgacccca aaaaacttga ttagggtgat
3401  ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac
3451  gttggagtcc acgttcttta atagtggact cttgttccaa actggaacaa
3501  cactcaaccc tatctcggtc tattcttttg atttataagg gattttgccg
3551  atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc
3601  gaattttaac aaaatattaa cgtttacaat ttcctgatgc ggtattttct
3651  ccttacgcat ctgtgcggta tttcacaccg catatggtgc actctcagta
3701  caatctgctc tgatgccgca tagttaagcc agccccgaca cccgccaaca
3751  cccgctgacg cgccctgacg ggcttgtctg ctcccggcat ccgcttacag
3801  acaagctgtg accgtctccg ggagctgcat gtgtcagagg ttttcaccgt
3851  catcaccgaa acgcgcgaga cgaaagggcc tcgtgatacg cctatttta
3901  taggttaatg tcatgataat aatggtttct tagacgtcag gtggcacttt
3951  tcggggaaat gtgcgcggaa cccctatttg tttatttttc taaatacatt
4001  caaatatgta tccgctcatg agacaataac cctgataaat gcttcaataa
4051  tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat
4101  tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc
4151  tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac
4201  atcgaactgg atctcaacag cggtaagatc cttgagagtt ttcgccccga
4251  agaacgtttt ccaatgatga gcacttttaa agttctgcta tgtggcgcgg
4301  tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac
4351  tattctcaga atgacttggt tgagtactca ccagtcacag aaaagcatct
4401  tacggatggc atgacagtaa gagaattatg cagtgctgcc ataaccatga
4451  gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag
4501  gagctaaccg ctttttgca caacatgggg gatcatgtaa ctcgccttga
4551  tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca
4601  ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc
4651  gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc
4701  ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggt
4751  ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt
4801  gcagcactgg ggccagatgg taagccctcc cgtatcgtag ttatctacac
4851  gacggggagt caggcaacta tggatgaacg aaatagacag atcgctgaga
4901  taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca
4951  tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta
5001  ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt
5051  tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct
5101  tgagatcctt ttttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc
5151  accgctacca gcggtggttt gtttgccgga tcaagagcta ccaactcttt
5201  ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgtcctt
5251  ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc
5301  tacatacctc gctctgctaa tcctgttacc agtggctgct gccagtggcg
5351  ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag
5401  gcgcagcggt cgggctgaac gggggttcg tgcacacagc ccagcttgga
5451  gcgaacgacc tacaccgaac tgagatacct acagcgtgag cattgagaaa
5501  gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc
5551  agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg
5601  gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat
5651  ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac
5701  gcggcctttt tacggttcct ggccttttgc tggccttttg ctcacatgtt
```

FIGURE 18B

```
5751  ctttcctgcg ttatcccctg attctgtgga taaccgtatt accgcctttg
5801  agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca
5851  gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc ctctccccgc
5901  gcgttggccg attcattaat gcagggctgc ag
```

FIGURE 18C

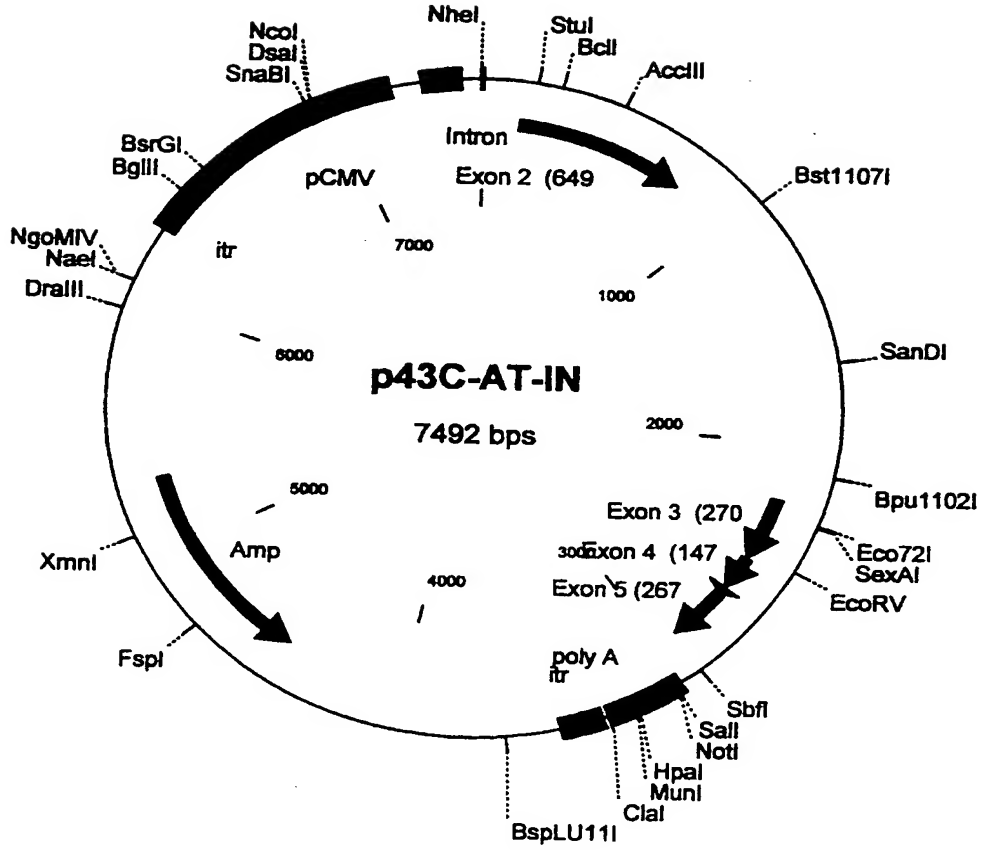

NcoI
DsaI
SnaBI
NheI
StuI
BclI
AccIII
BsrGI
BglII
Intron
pCMV
Exon 2 (649
Bst1107I
NgoMIV
NaeI
DraIII
itr
7000
1000
6000
p43C-AT-IN
7492 bps
SanDI
2000
5000
Exon 3 (270
Bpu1102I
Amp
Exon 4 (147
Eco72I
SexAI
XmnI
4000
Exon 5 (267
EcoRV
poly A
itr
FspI
SbfI
SalI
NotI
HpaI
MunI
ClaI
BspLU11I

FIGURE 19

Molecule Name: p43C-AT-IN 7492 bps DNA Circular
Sequence Printed: 1-7492 (Full) Date Printed 16 Apr 1999
Description: Ligation of p43-C into IN

```
   1  aattccgcat tgcagagata attgtattta agtgcctagc tcgatacaat
  51  aaacgccatt tgaccattca ccacattggt gtgcacctcc actagctgcc
 101  ttgactgcct ggccccccca tctctgtctt gcaggacaat gccgtcttct
 151  gtctcgtggg gcatcctcct gctggcaggc ctgtgctgcc tggtccctgt
 201  ctccctggct gaggatcccc agggagatgc tgcccagaag acagatacat
 251  cccaccatga tcaggatcac ccaaccttca acaagatcac ccccaacctg
 301  gctgagttcg ccttcagcct ataccgccag ctggcacacc agtccaacag
 351  caccaatatc ttcttctccc cagtgagcat cgctacagcc tttgcaatgc
 401  tctccctggg gaccaaggct gacactcacg atgaaatcct ggagggcctg
 451  aatttcaacc tcacggagat tccggaggct cagatccatg aaggcttcca
 501  ggaactcctc cgtaccctca accagccaga cagccagctc cagctgacca
 551  ccggcaatgg cctgttcctc agcgagggcc tgaagctagt ggataagttt
 601  ttggaggatg ttaaaaagtt gtaccactca gaagccttca ctgtcaactt
 651  cggggacacc gaagaggcca agaaacagat caacgattac gtggagaagg
 701  gtactcaagg gaaaattgtg gatttggtca aggagcttga cagagacaca
 751  gtttttgctc tggtgaatta catcttcttt aaaggtaagg ttgctcaacc
 801  agcctgagct gtttcccata gaaacaagca aaaatatttc tcaaaccatc
 851  agttcttgaa ctctccttgg caatgcatta tgggccatag caatgctttt
 901  cagcgtggat tcttcagttt tctacacaca aacactaaaa tgttttccat
 951  cattgagtaa tttgaggaaa taatagatta aactgtcaaa actactgacg
1001  ctctgcagaa cttttcagag cctttaatgt ccttgtgtat actgtatatg
1051  tagaatatat aatgcttaga actatagaac aaattgtaat acactgcata
1101  aagggatagt ttcatggaac atactttaca cgactctagt gtcccagaat
1151  cagtatcagt tttgcaatct gaaagacctg ggttcaaatc ctgcctctaa
1201  cacaattagc ttttgacaaa aacaatgcat tctacctctt tgaggtgcta
1251  atttctcatc ttagcatgga caaaatacca ttcttgctgt caggtttttt
1301  taggattaaa caaatgacaa agactgtggg gatggtgtgt ggcatacagc
1351  aggtgatgga ctcttctgta tctcaggctg ccttcctgcc cctgaggggt
1401  taaaatgcca gggtcctggg ggccccaggg cattctaagc cagctcccac
1451  tgtcccagga aaacagcata ggggagggga ggtgggaggc aaggccaggg
1501  gctgcttcct ccactctgag gctcccttgc tcttgaggca aaggagggca
1551  gtggaggcaa gccaggctgc agtcagcaca gctaaagtcc tggctctgct
1601  gtggccttag tgggggccca ggtccctctc cagccccagt ctcctccttc
1651  tgtccaatga gaaagctggg atcaggggtc cctgaggccc ctgtccactc
1701  tgcatgcctc gatggtgaag ctctgttggt atggcagagg ggaggctgct
1751  caggcatctg catttcccct gccaatctag aggatgagga aagctctcag
1801  gaatagtaag cagaatgttt gccctggatg aataactgag ctgccaatta
1851  acaaggggca gggagcctta gacagaaggt accaaatatg cctgatgctc
1901  caacatttta tttgtaatat ccaagacacc ctcaaataaa catatgattc
1951  caataaaaat gcacagccac gatggcatct cttagcctga catcgccacg
2001  atgtagaaat tctgcatctt cctctagttt tgaattatcc ccacacaatc
2051  tttttcggca gcttggatgg tcagtttcag caccttttac agatgatgaa
2101  gctgagcctc gagggatgtg tgtcgtcaag ggggctcagg gcttctcagg
2151  gaggggactc atggtttctt attctgctac actcttccaa accttcactc
2201  acccctggtg atgcccacct tcccctctct ccaggcaaat gggagagacc
2251  ctttgaagtc aaggacaccg aggaagagga cttccacgtg gaccaggtga
2301  ccaccgtgaa ggtgcctatg atgaagcgtt taggcatgtt taacatccag
2351  cactgtaaga agctgtccag ctgggtgctg ctgatgaaat acctgggcaa
2401  tgccaccgcc atcttcttcc tgcctgatga ggggaaacta cagcacctgg
2451  aaaatgaact cacccacgat atcatcacca agttcctgga aaatgaagac
2501  agaaggtctg ccagcttaca tttacccaaa ctgtccatta ctggaaccta
2551  tgatctgaag agcgtcctgg gtcaactggg catcactaag gtcttcagca
2601  atggggctga cctctccggg gtcacagagg aggcacccct gaagctctcc
2651  aaggccgtgc ataaggctgt gctgaccatc gacgagaaag ggactgaagc
2701  tgctggggcc atgtttttag aggccatacc catgtctatc ccccccgagg
```

FIGURE 19A

```
2751  tcaagttcaa caaacccttt gtcttcttaa tgattgaaca aaataccaag
2801  tctccccctct tcatgggaaa agtggtgaat cccacccaaa aataactgcc
2851  tctcgctcct caacccctcc cctccatccc tggcccctc cctggatgac
2901  attaaagaag ggttgagctg gtaacccccc ccccccctgc aggccctcga
2951  gacgcgtggc atgcaagctt ggtaccgagc tcggatccac tagtaacggc
3001  cgccagtgtg ctggaattca cgcgtggtac ctctagagtc gacccgggcg
3051  gccgcttcga gcagacatga taagatacat tgatgagttt ggacaaacca
3101  caactagaat gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct
3151  attgctttat ttgtaaccat tataagctgc aataaacaag ttaacaacaa
3201  caattgcatt cattttatgt ttcaggttca gggggagatg tgggaggttt
3251  tttaaagcaa gtaaaacctc tacaaatgtg gtaaaatcga taaggatcta
3301  ggaaccccta gtgatggagt tggccactcc ctctctgcgc gctcgctcgc
3351  tcactgaggc cgcccgggca aagcccgggc gtcgggcgac ctttggtcgc
3401  ccggcctcag tgagcgagcg agcgcgcaga gagggagtgg ccaactgcag
3451  ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg
3501  gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc
3551  tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca
3601  gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa
3651  ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc
3701  gcccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga
3751  aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct
3801  cgtgcgctct cctgttccga ccctgccgct taccggatac ctgtccgcct
3851  ttctcccttc gggaagcgtg gcgctttctc aatgctcacg ctgtaggtat
3901  ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc
3951  ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt
4001  ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac
4051  aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg
4101  gtggcctaac tacggctaca ctagaaggac agtatttggt atctgcgctc
4151  tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc
4201  aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat
4251  tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg
4301  ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg
4351  agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag
4401  ttttaaatca atctaaagta tatatgagta aacttggtct gacagttacc
4451  aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca
4501  tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg
4551  cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac
4601  cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc
4651  agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg
4701  ccgggaagct agagtaagta gttcgccagt taatagtttg cgcaacgttg
4751  ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct
4801  tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat
4851  gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc gttgtcagaa
4901  gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat
4951  tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta
5001  ctcaaccaag tcattctgag aatagtgtat gcggcgaccg agttgctctt
5051  gcccggcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa
5101  gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt
5151  accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat
5201  cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga
5251  aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat
5301  actcatactc ttcctttttc aatattattg aagcatttat cagggttatt
5351  gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata
5401  ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac
5451  cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct
5501  ttcgtctcgc gcgtttcggt gatgacggtg aaaacctctg acacatgcag
5551  ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca
5601  agcccgtcag ggcgcgtcag cgggtgttgg cgggtgtcgg ggctggctta
5651  actatgcggc atcagagcag attgtactga gagtgcacca tatgcggtgt
5701  gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggaaattgta
```

FIGURE 19B

```
5751  aacgttaata ttttgttaaa attcgcgtta aatttttgtt aaatcagctc
5801  atttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag
5851  aatagaccga gatagggttg agtgttgttc cagtttggaa caagagtcca
5901  ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca
5951  gggcgatggc ccactacgtg aaccatcacc ctaatcaagt tttttggggt
6001  cgaggtgccg taaagcacta aatcggaacc ctaaagggag cccccgattt
6051  agagcttgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa
6101  agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc
6151  gcgtaaccac cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcg
6201  cgccattcgc cattcaggct acgcaactgt tgggaagggc gatcggtgcg
6251  ggcctcttcg ctattacgcc agctgcagtt ggccactccc tctctgcgcg
6301  ctcgctcgct cactgaggcc gggcgaccaa aggtcgcccg acgcccgggc
6351  tttgcccggg cggcctcagt gagcgagcga gcgcgcagag agggagtggc
6401  caactccatc actaggggtt cctagatctt caatattggc cattagccat
6451  attattcatt ggttatatag cataaatcaa tattggctat tggccattgc
6501  atacgttgta tctatatcat aatatgtaca tttatattgg ctcatgtcca
6551  atatgaccgc catgttggca ttgattattg actagttatt aatagtaatc
6601  aattacgggg tcattagttc atagcccata tatggagttc cgcgttacat
6651  aacttacggt aaatggcccg cctggctgac cgcccaacga cccccgccca
6701  ttgacgtcaa taatgacgta tgttcccata gtaacgccaa tagggacttt
6751  ccattgacgt caatgggtgg agtatttacg gtaaactgcc cacttggcag
6801  tacatcaagt gtatcatatg ccaagtccgc cccctattga cgtcaatgac
6851  ggtaaatggc ccgcctggca ttatgcccag tacatgacct tacgggactt
6901  tcctacttgg cagtacatct acgtattagt catcgctatt accatggtga
6951  tgcggttttg gcagtacacc aatgggcgtg gatagcggtt tgactcacgg
7001  ggatttccaa gtctccaccc cattgacgtc aatgggagtt tgttttggca
7051  ccaaaatcaa cgggactttc caaaatgtcg taataacccc gccccgttga
7101  cgcaaatggg cggtaggcgt gtacggtggg aggtctatat aagcagagct
7151  cgtttagtga accgtcagat cactagaagc tttattgcgg tagtttatca
7201  cagttaaatt gctaacgcag tcagtgcttc tgacacaaca gtctcgaact
7251  taagctgcag aagttggtcg tgaggcactg ggcaggtaag tatcaaggtt
7301  acaagacagg tttaaggaga ccaatagaaa ctgggcttgt cgagacagag
7351  aagactcttg cgtttctgat aggcacctat tggtcttact gacatccact
7401  ttgcctttct ctccacaggt gtccactccc agttcaatta cagctcttaa
7451  ggctagagta cttaatacga ctcactatag gctagcctcg ag
```

FIGURE 19C

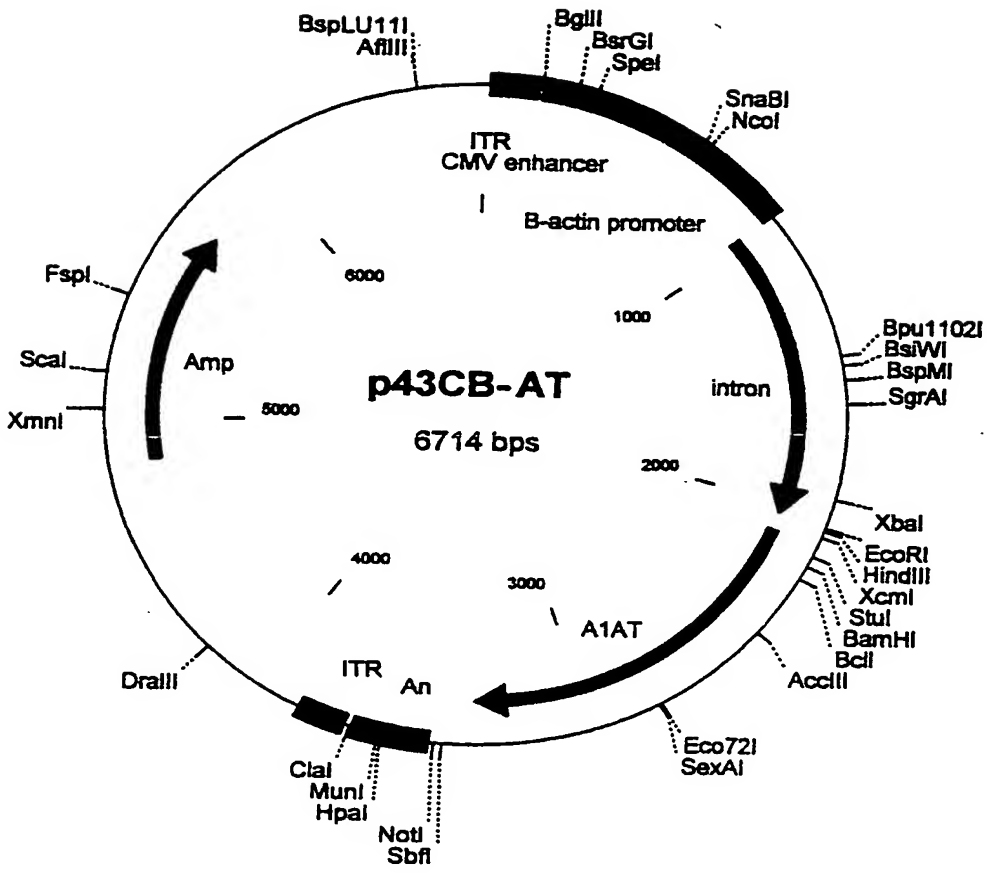

BspLU11I
AflIII
BglII
BsrGI
SpeI
SnaBI
NcoI
ITR
CMV enhancer
B-actin promoter
1000
Bpu1102I
BsiWI
BspMI
SgrAI
intron
2000
XbaI
EcoRI
HindIII
XcmI
StuI
BamHI
BclI
AccIII
3000
A1AT
Eco72I
SexAI
4000
ITR
An
ClaI
MunI
HpaI
NotI
SbfI
DraIII
5000
Amp
6000
FspI
ScaI
XmnI
p43CB-AT
6714 bps

FIGURE 20

16 Apr 1999 Sequence Data Page 1

Molecule: p43CB-AT, 6714 bps DNA Circular
Description: Ligation of Fragment 2 into Fragment 2
File Name: CB-AAT.cm5, dated 17 Nov 1998
Printed: 1-6714 bps (Full), format Single Strand

```
   1  gggggggggg gggggggttg gccactccct ctctgcgcgc tcgctcgctc
  51  actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc
-101  ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca
 151  ctaggggttc ctagatcttc aatattggcc attagccata ttattcattg
 201  gttatatagc ataaatcaat attggctatt ggccattgca tacgttgtat
 251  ctatatcata atatgtacat ttatattggc tcatgtccaa tatgaccgcc
 301  atgttggcat tgattattga ctagttatta atagtaatca attacggggt
 351  cattagttca tagcccatat atggagttcc gcgttacata acttacggta
 401  aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat
 451  aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc
 501  aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg
 551  tatcatatgc caagtccgcc ccctattgac gtcaatgacg gtaaatggcc
 601  cgcctggcat tatgcccagt acatgacctt acgggacttt cctacttggc
 651  agtacatcta cgtattagtc atcgctatta ccatggtcga ggtgagcccc
 701  acgttctgct tcactctccc catctccccc ccctccccac ccccaatttt
 751  gtatttattt atttttaat tattttgtgc agcgatgggg gcgggggggg
 801  ggggggggcg cgcgccaggc ggggcggggc ggggcgaggg gcggggcggg
 851  gcgaggcgga gaggtgcggc ggcagccaat cagagcggcg cgctccgaaa
 901  gtttcctttt atggcgaggc ggcggcggcg gcggccctat aaaaagcgaa
 951  gcgcgcggcg ggcgggagtc gctgcgacgc tgccttcgcc ccgtgccccg
1001  ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg accgcgttac
1051  tcccacaggt gagcgggcgg gacggcccct ctcctccggg ctgtaattag
1101  cgcttggttt aatgacggct tgtttctttt ctgtggctgc gtgaaagcct
1151  tgaggggctc cgggagggcc ctttgtgcgg gggggagcgg ctcggggggt
1201  gcgtgcgtgt gtgtgtgcgt ggggagcgcc gcgtgcggcc cgcgctgccc
1251  ggcggctgtg agcgctgcgg gcgcggcgcg gggctttgtg cgctccgcag
1301  tgtgcgcgag gggagcgcgg ccgggggcgg tgccccgcgg tgcgggggg
1351  gctgcgaggg gaacaaaggc tgcgtgcggg gtgtgtgcgt gggggggtga
1401  gcaggggggtg tgggcgcggc ggtcgggctg taacccccc ctgcacccc
1451  ctccccgagt tgctgagcac ggcccggctt cgggtgcggg gctccgtacg
1501  gggcgtggcg cggggctcgc cgtgccgggc gggggggtggc ggcaggtggg
1551  ggtgccgggc ggggcggggc cgcctcgggc cggggagggc tcgggggagg
1601  ggcgcggcgg cccccggagc gccggcggct gtcgaggcgc ggcgagccgc
1651  agccattgcc ttttatggta atcgtgcgag agggcgcagg gacttccttt
1701  gtcccaaatc tgtgcggagc cgaaatctgg gaggcgccgc cgcacccccct
1751  ctagcgggcg cggggcgaag cggtgcggcg ccggcaggaa ggaaatgggc
1801  ggggagggcc ttcgtgcgtc gccgcgccgc cgtcccctc tccctctcca
1851  gcctcggggc tgtccgcggg gggacggctg ccttcggggg ggacggggca
1901  gggcggggtt cggcttctgg cgtgtgaccg gcggctctag agcctctgct
1951  aaccatgttc atgccttctt ctttttccta cagctcctgg gcaacgtgct
2001  ggttattgtg ctgtctcatc attttggcaa agaattcgat atcaagcttg
2051  gggattttca ggcaccacca ctgacctggg acagtgaatc gacaatgccg
2101  tcttctgtct cgtggggcat cctcctgctg gcaggcctgt gctgcctggt
2151  ccctgtctcc ctggctgagg atccccaggg agatgctgcc cagaagacag
2201  atacatccca ccatgatcag gatcacccaa ccttcaacaa gatcaccccc
2251  aacctggctg agttcgcctt cagcctatac cgccagctgg cacaccagtc
2301  caacagcacc aatatcttct tctccccagt gagcatcgct acagcctttg
2351  caatgctctc cctggggacc aaggctgaca ctcacgatga aatcctggag
2401  ggcctgaatt tcaacctcac ggagattccg gaggctcaga tccatgaagg
```

FIGURE 20A

```
2451  cttccaggaa ctcctccgta ccctcaacca gccagacagc cagctccagc
2501  tgaccaccgg caatggcctg ttcctcagcg agggcctgaa gctagtggat
2551  aagtttttgg aggatgttaa aaagttgtac cactcagaag ccttcactgt
2601  caacttcggg gacaccgaag aggccaagaa acagatcaac gattacgtgg
2651  agaagggtac tcaagggaaa attgtggatt tggtcaagga gcttgacaga
2701  gacacagttt ttgctctggt gaattacatc ttctttaaag gcaaatggga
2751  gagacccttt gaagtcaagg acaccgagga agaggacttc cacgtggacc
2801  aggtgaccac cgtgaaggtg cctatgatga agcgtttagg catgtttaac
2851  atccagcact gtaagaagct gtccagctgg gtgctgctga tgaaatacct
2901  gggcaatgcc accgccatct tcttcctgcc tgatgagggg aaactacagc
2951  acctggaaaa tgaactcacc cacgatatca tcaccaagtt cctggaaaat
3001  gaagacagaa ggtctgccag cttacattta cccaaactgt ccattactgg
3051  aacctatgat ctgaagagcg tcctgggtca actgggcatc actaaggtct
3101  tcagcaatgg ggctgacctc tccggggtca cagaggaggc acccctgaag
3151  ctctccaagg ccgtgcataa ggctgtgctg accatcgacg agaaagggac
3201  tgaagctgct ggggccatgt ttttagaggc catacccatg tctatccccc
3251  ccgaggtcaa gttcaacaaa ccctttgtct tcttaatgat tgaacaaaat
3301  accaagtctc ccctcttcat gggaaaagtg gtgaatccca cccaaaaata
3351  actgcctctc gctcctcaac ccctcccctc catccctggc cccctccctg
3401  gatgacatta aagaagggtt gagctggtaa ccccccccccc ccctgcaggg
3451  gccctcgacc cgggcggccg cttcgagcag acatgataag atacattgat
3501  gagtttggac aaaccacaac tagaatgcag tgaaaaaaat gctttatttg
3551  tgaaatttgt gatgctattg ctttatttgt aaccattata agctgcaata
3601  aacaagttaa caacaacaat tgcattcatt ttatgtttca ggttcagggg
3651  gagatgtggg aggtttttta aagcaagtaa aacctctaca aatgtggtaa
3701  aatcgataag gatctaggaa cccctagtga tggagttggc cactccctct
3751  ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg
3801  ggcgaccttt ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg
3851  gagtggccaa ccccccccccc ccccccctg cagcctggcg taatagcgaa
3901  gaggcccgca ccgatcgccc ttcccaacag ttgcgtagcc tgaatggcga
3951  atggcgcgac gcgccctgta gcggcgcatt aagcgcggcg ggtgtggtgg
4001  ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct
4051  ttcgctttct tcccttcctt tctcgccacg ttcgccggct ttccccgtca
4101  agctctaaat cgggggctcc ctttagggtt ccgatttagt gctttacggc
4151  acctcgaccc caaaaaactt gattagggtg atggttcacg tagtgggcca
4201  tcgccctgat agacggtttt tcgccctttg acgttggagt ccacgttctt
4251  taatagtgga ctcttgttcc aaactggaac aacactcaac cctatctcgg
4301  tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta
4351  aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt
4401  aacgtttaca atttcctgat gcggtatttt ctccttacgc atctgtgcgg
4451  tatttcacac cgcatatggt gcactctcag tacaatctgc tctgatgccg
4501  catagttaag ccagccccga cacccgccaa cacccgctga cgcgccctga
4551  cgggcttgtc tgctcccggc atccgcttac agacaagctg tgaccgtctc
4601  cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg aaacgcgcga
4651  gacgaaaggg cctcgtgata cgcctatttt tataggttaa tgtcatgata
4701  ataatggttt cttagacgtc aggtggcact tttcggggaa atgtgcgcgg
4751  aacccctatt tgtttatttt tctaaataca ttcaaatatg tatccgctca
4801  tgagacaata accctgataa atgcttcaat aatattgaaa aaggaagagt
4851  atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt
4901  ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg
4951  ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac
5001  agcggtaaga tccttgagag ttttcgcccc gaagaacgtt ttccaatgat
5051  gagcactttt aaagttctgc tatgtggcgc ggtattatcc cgtattgacg
5101  ccgggcaaga gcaactcggt cgccgcatac actattctca gaatgacttg
5151  gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt
```

FIGURE 20B

```
5201  aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca
5251  acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg
5301  cacaacatgg gggatcatgt aactcgcctt gatcgttggg aaccggagct
5351  gaatgaagcc ataccaaacg acgagcgtga caccacgatg cctgtagcaa
5401  tggcaacaac gttgcgcaaa ctattaactg gcgaactact tactctagct
5451  tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc
5501  acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg
5551  gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat
5601  ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac
5651  tatggatgaa cgaaatagac agatcgctga gataggtgcc tcactgatta
5701  agcattggta actgtcagac caagtttact catatatact ttagattgat
5751  ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttttga
5801  taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt
5851  cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttttctg
5901  cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt
5951  ttgtttgccg gatcaagagc taccaactct ttttccgaag gtaactggct
6001  tcagcagagc gcagatacca aatactgtcc ttctagtgta gccgtagtta
6051  ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct
6101  aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg
6151  ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga
6201  acgggggggtt cgtgcacaca gcccagcttg gagcgaacga cctacaccga
6251  actgagatac ctacagcgtg agcattgaga aagcgccacg cttcccgaag
6301  ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag
6351  cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt
6401  cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag
6451  gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc
6501  ctggcctttt gctggccttt tgctcacatg ttctttcctg cgttatcccc
6551  tgattctgtg gataaccgta ttaccgcctt tgagtgagct gataccgctc
6601  gccgcagccg aacgaccgag cgcagcgagt cagtgagcga ggaagcggaa
6651  gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta
6701  atgcagggct gcag
```

FIGURE 20C

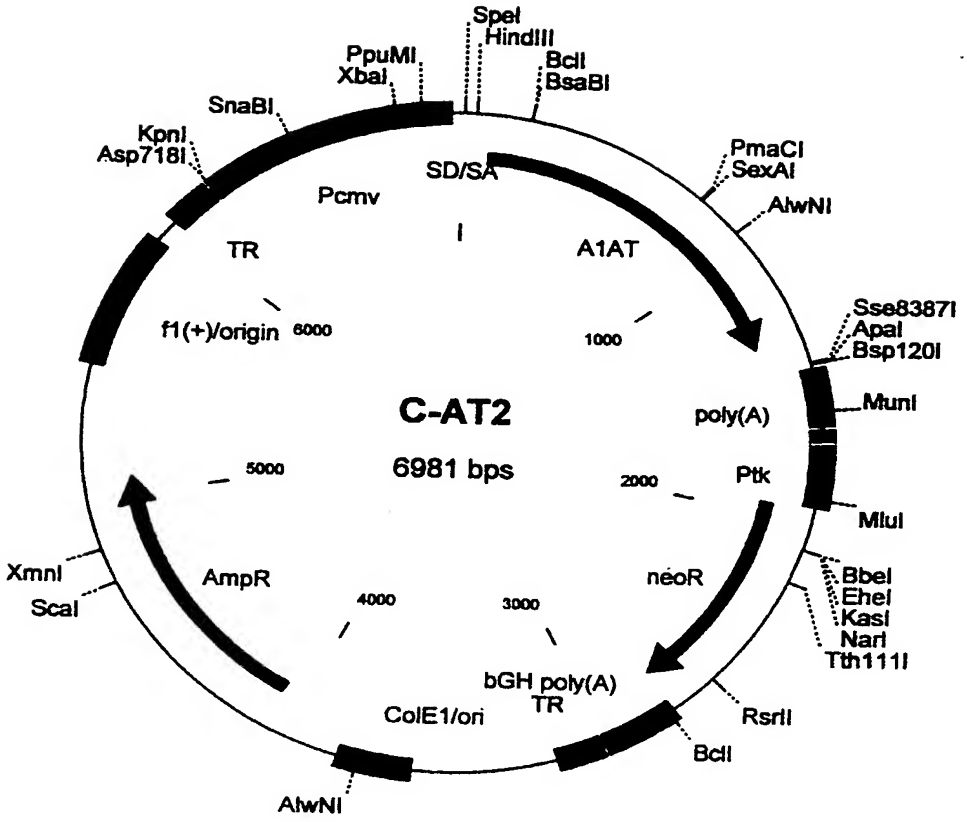

SpeI
HindIII
BclI
BsaBI
PpuMI
XbaI
SnaBI
KpnI
Asp718I
PmaCI
SexAI
AlwNI
SD/SA
Pcmv
TR
A1AT
f1(+)/origin
6000
1000
Sse8387I
ApaI
Bsp120I
MunI
C-AT2
6981 bps
poly(A)
5000
2000
Ptk
MluI
XmnI
ScaI
AmpR
neoR
BbeI
EheI
KasI
NarI
Tth111I
4000
3000
bGH poly(A)
TR
ColE1/ori
RsrII
BclI
AlwNI

FIGURE 21

Molecule Name: C-AT2 6981 bps DNA Circular
Sequence Printed: 1-6981 (Full) Date Printed 16 Apr 1999
Description: Ligation of Fragment 1 and Fragment 2

```
   1  ctagaactag tggatccccc gggctgcagg aattcgatat caagcttggg
  51  gattttcagg caccaccact gacctgggac agtgaatcga caatgccgtc
 101  ttctgtctcg tggggcatcc tcctgctggc aggcctgtgc tgcctggtcc
 151  ctgtctccct ggctgaggat ccccagggag atgctgccca gaagacagat
 201  acatcccacc atgatcagga tcacccaacc ttcaacaaga tcacccccaa
 251  cctggctgag ttcgccttca gcctataccg ccagctggca caccagtcca
 301  acagcaccaa tatcttcttc tccccagtga gcatcgctac agcctttgca
 351  atgctctccc tggggaccaa ggctgacact cacgatgaaa tcctggaggg
 401  cctgaatttc aacctcacgg agattccgga ggctcagatc catgaaggct
 451  tccaggaact cctccgtacc ctcaaccagc cagacagcca gctccagctg
 501  accaccggca atggcctgtt cctcagcgag ggcctgaagc tagtggataa
 551  gtttttggag gatgttaaaa agttgtacca ctcagaagcc ttcactgtca
 601  acttcgggga caccgaagag gccaagaaac agatcaacga ttacgtggag
 651  aagggtactc aagggaaaat tgtggatttg gtcaaggagc ttgacagaga
 701  cacagttttt gctctggtga attacatctt ctttaaaggc aaatgggaga
 751  gacccttga agtcaaggac accgaggaag aggacttcca cgtggaccag
 801  gtgaccaccg tgaaggtgcc tatgatgaag cgtttaggca tgtttaacat
 851  ccagcactgt aagaagctgt ccagctgggt gctgctgatg aaatacctgg
 901  gcaatgccac cgccatcttc ttcctgcctg atgaggggaa actacagcac
 951  ctggaaaatg aactcaccca cgatatcatc accaagttcc tggaaaatga
1001  agacagaagg tctgccagct tacatttacc caaactgtcc attactggaa
1051  cctatgatct gaagagcgtc ctgggtcaac tgggcatcac taaggtcttc
1101  agcaatgggg ctgacctctc cggggtcaca gaggaggcac ccctgaagct
1151  ctccaaggcc gtgcataagg ctgtgctgac catcgacgag aaagggactg
1201  aagctgctgg ggccatgttt ttagaggcca tacccatgtc tatccccccc
1251  gaggtcaagt tcaacaaacc ctttgtcttc ttaatgattg aacaaaatac
1301  caagtctccc ctcttcatgg gaaaagtggt gaatcccacc caaaaataac
1351  tgcctctcgc tcctcaaccc ctcccctcca tccctggccc cctccctgga
1401  tgacattaaa gaagggttga gctggtaacc ccccccccc ctgcaggggc
1451  cctcgaggcc gcggggatcc agacatgata agatacattg atgagtttgg
1501  acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt
1551  gtgatgctat tgctttattt gtaaccatta taagctgcaa taaacaagtt
1601  aacaacaaca attgcattca ttttatgttt caggttcagg gggaggtgtg
1651  ggaggttttt tagtcgacct cgagcagtgt ggttttgcaa gaggaagcaa
1701  aaagcctctc cacccaggcc tggaatgttt ccacccaagt cgaaggcagt
1751  gtggttttgc aagaggaagc aaaaagcctc tccacccagg cctggaatgt
1801  ttccacccaa tgtcgagcaa ccccgcccag cgtcttgtca ttggcgaatt
1851  cgaacacgca gatgcagtcg gggcggcgcg gtcccaggtc cacttcgcat
1901  attaaggtga cgcgtgtggc ctcgaacacc gagcgaccct gcagccaata
1951  tgggatcggc cattgaacaa gatggattgc acgcaggttc tccggccgct
2001  tgggtggaga ggctattcgg ctatgactgg gcacaacaga caatcggctg
2051  ctctgatgcc gccgtgttcc ggctgtcagc gcaggggcgc ccggttcttt
2101  ttgtcaagac cgacctgtcc ggtgccctga atgaactgca ggacgaggca
2151  gcgcggctat cgtggctggc cacgacgggc gttccttgcg cagctgtgct
2201  cgacgttgtc actgaagcgg gaagggactg gctgctattg ggcgaagtgc
2251  cggggcagga tctcctgtca tctcaccttg ctcctgccga gaaagtatcc
2301  atcatggctg atgcaatgcg gcggctgcat acgcttgatc cggctacctg
2351  cccattcgac caccaagcga aacatcgcat cgagcgagca cgtactcgga
2401  tggaagccgg tcttgtcgat caggatgatc tggacgaaga gcatcagggg
2451  ctcgcgccag ccgaactgtt cgccaggctc aaggcgcgca tgcccgacgg
2501  cgaggatctc gtcgtgaccc atggcgatgc ctgcttgccg aatatcatgg
2551  tggaaaatgg ccgcttttct ggattcatcg actgtggccg gctgggtgtg
2601  gcggaccgct atcaggacat agcgttggct acccgtgata ttgctgaaga
2651  gcttggcggc gaatgggctg accgcttcct cgtgctttac ggtatcgccg
2701  ctcccgattc gcagcgcatc gccttctatc gccttcttga cgagttcttc
```

FIGURE 21A

```
2751  tgaggggatc cgtcgactag agctcgctga tcagcctcga ctgtgccttc
2801  tagttgccag ccatctgttg tttgccctc ccccgtgcct tccttgaccc
2851  tggaaggtgc cactcccact gtcctttcct aataaaatga ggaaattgca
2901  tcgcattgtc tgagtaggtg tcattctatt ctgggggggtg gggtggggca
2951  ggacagcaag ggggaggatt gggaagacaa tagcaggcat gctggggaga
3001  gatctaggaa cccctagtga tggagttggc cactccctct ctgcgcgctc
3051  gctcgctcac tgaggccgcc cgggcaaagc ccgggcgtcg ggcgaccttt
3101  ggtcgcccgg cctcagtgag cgagcgagcg cgcagagagg gagtggccaa
3151  ccccccccc ccccccctg cagccctgca ttaatgaatc ggccaacgcg
3201  cggggagagg cggtttgcgt attgggcgct cttccgcttc ctcgctcact
3251  gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat cagctcactc
3301  aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga
3351  acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg
3401  ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa
3451  tcgacgctca agtcagaggt ggcgaaaccc gacaggacta taaagatacc
3501  aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt tccgaccctg
3551  ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct
3601  ttctcaatgc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct
3651  ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc
3701  ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc
3751  gccactggca gcagccactg gtaacaggat tagcagagcg aggtatgtag
3801  gcggtgctac agagttcttg aagtggtggc ctaactacgg ctacactaga
3851  aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa
3901  aagagttggt agctcttgat ccggcaaaca aaccaccgct ggtagcggtg
3951  gtttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa
4001  gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa
4051  ctcacgttaa gggattttgg tcatgagatt atcaaaaagg atcttcacct
4101  agatcctttt aaattaaaaa tgaagtttta aatcaatcta aagtatatat
4151  gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat
4201  ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccgtcg
4251  tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca
4301  atgataccgc gagacccacg ctcaccggct ccagatttat cagcaataaa
4351  ccagccagcc ggaagggccg agcgcagaag tggtcctgca actttatccg
4401  cctccatcca gtctattaat tgttgccggg aagctagagt aagtagttcg
4451  ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt
4501  gtcacgctcg tcgtttggta tggcttcatt cagctccggt tcccaacgat
4551  caaggcgagt tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc
4601  ttcggtcctc cgatcgttgt cagaagtaag ttggccgcag tgttatcact
4651  catggttatg gcagcactgc ataattctct tactgtcatg ccatccgtaa
4701  gatgcttttc tgtgactggt gagtactcaa ccaagtcatt ctgagaatag
4751  tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac
4801  cgcgccacat agcagaactt taaaagtgct catcattgga aaacgttctt
4851  cggggcgaaa actctcaagg atcttaccgc tgttgagatc cagttcgatg
4901  taacccactc gtgcacccaa ctgatcttca gcatctttta ctttcaccag
4951  cgtttctggg tgagcaaaaa caggaaggca aaatgccgca aaaaagggaa
5001  taagggcgac acggaaatgt tgaatactca tactcttcct ttttcaatat
5051  tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga
5101  atgtatttag aaaaataaac aaatagggggt tccgcgcaca tttccccgaa
5151  aagtgccacc tgacgtctaa gaaaccatta ttatcatgac attaacctat
5201  aaaaataggc gtatcacgag gccctttcgt ctcgcgcgtt tcggtgatga
5251  cggtgaaaac ctctgacaca tgcagctccc ggagacggtc acagcttgtc
5301  tgtaagcgga tgccgggagc agacaagccc gtcagggcgc gtcagcgggt
5351  gttggcgggt gtcggggctg gcttaactat gcggcatcag agcagattgt
5401  actgagagtg caccatatgc ggtgtgaaat accgcacaga tgcgtaagga
5451  gaaaataccg catcaggaaa ttgtaaacgt taatattttg ttaaaattcg
5501  cgttaaattt ttgttaaatc agctcatttt ttaaccaata ggccgaaatc
5551  ggcaaaatcc cttataaatc aaaagaatag accgagatag ggttgagtgt
5601  tgttccagtt tggaacaaga gtccactatt aaagaacgtg gactccaacg
5651  tcaaagggcg aaaaaccgtc tatcagggcg atggcccact acgtgaacca
5701  tcaccctaat caagtttttt ggggtcgagg tgccgtaaag cactaaatcg
```

FIGURE 21B

```
5751  gaaccctaaa gggagccccc gatttagagc ttgacgggga aagccggcga
5801  acgtggcgag aaaggaaggg aagaaagcga aaggagcggg cgctagggcg
5851  ctggcaagtg tagcggtcac gctgcgcgta accaccacac ccgccgcgct
5901  taatgcgccg ctacagggcg cgtcgcgcca ttcgccattc aggctacgca
5951  actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccaggct
6001  gcagggggggg ggggggggggg gttggccact ccctctctgc gcgctcgctc
6051  gctcactgag gccgggcgac caaaggtcgc ccgacgcccg ggctttgccc
6101  gggcggcctc agtgagcgag cgagcgcgca gagagggagt ggccaactcc
6151  atcactaggg gttcctagat ctgaattcgg tacccgttac ataacttacg
6201  gtaaatggcc cgcctggctg accgcccaac gacccccgcc cattgacgtc
6251  aataatgacg tatgttccca tagtaacgcc aatagggact ttccattgac
6301  gtcaatgggt ggagtattta cggtaaactg cccacttggc agtacatcaa
6351  gtgtatcata tgccaagtac gccccctatt gacgtcaatg acggtaaatg
6401  gcccgcctgg cattatgccc agtacatgac cttatgggac tttcctactt
6451  ggcagtacat ctacgtatta gtcatcgcta ttaccatggt gatgcggttt
6501  tggcagtaca tcaatgggcg tggatagcgg tttgactcac ggggatttcc
6551  aagtctccac cccattgacg tcaatgggag tttgttttgg caccaaaatc
6601  aacgggactt tccaaaatgt cgtaacaact ccgccccatt gacgcaaatg
6651  ggcggtaggc gtgtacggtg ggaggtctat ataagcagag ctcgtttagt
6701  gaaccgtcag atcgcctgga gacgccatcc acgctgtttt gacctccata
6751  gaagacaccg ggaccgatcc agcctccgga ctctagagga tccggtactc
6801  gaggaactga aaaaccagaa agttaactgg taagtttagt ctttttgtct
6851  tttatttcag gtcccggatc cggtggtggt gcaaatcaaa gaactgctcc
6901  tcagtggatg ttgcctttac ttctaggcct gtacggaagt gttacttctg
6951  ctctaaaagc tgcggaattg tacccgcggc c
```

FIGURE 21C

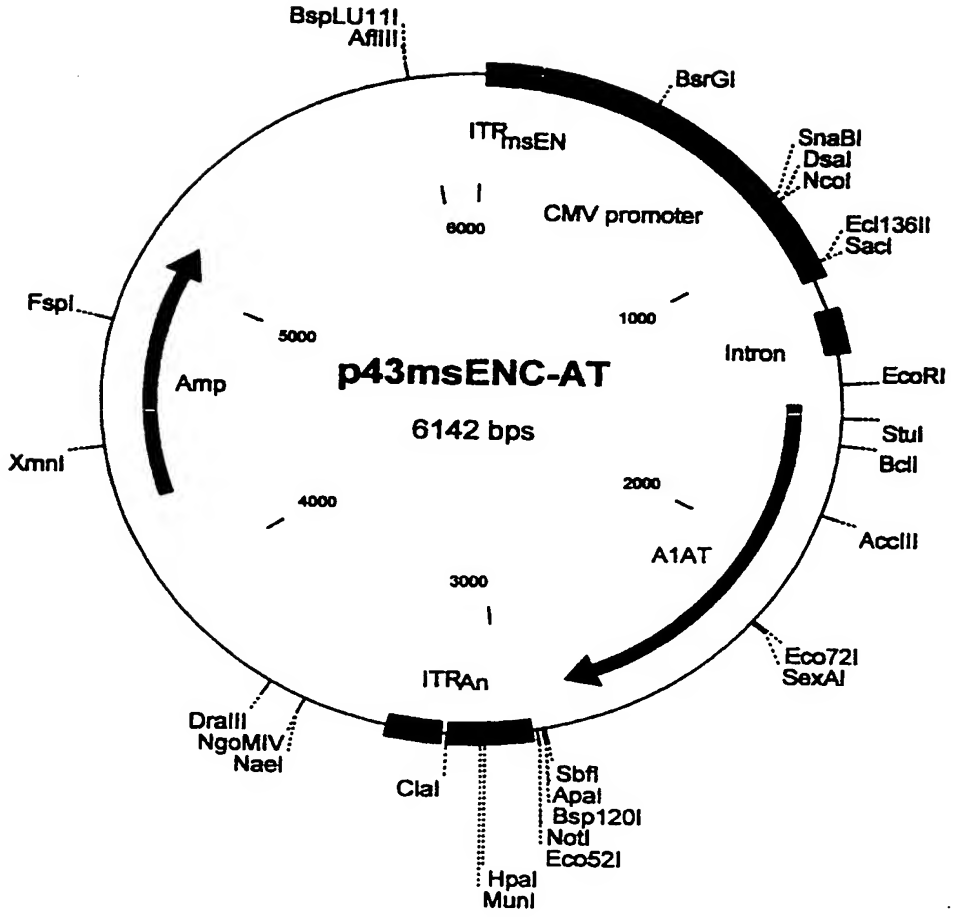

BspLU11I
AflIII
BsrGI
ITRmsEN
SnaBI
DsaI
NcoI
CMV promoter
6000
Ecl136II
SacI
FspI
1000
5000
Intron
p43msENC-AT
Amp
EcoRI
6142 bps
StuI
BclI
Xmnl
2000
4000
AccIII
A1AT
3000
Eco72I
SexAI
ITRAn
DraIII
NgoMIV
NaeI
ClaI
SbfI
ApaI
Bsp120I
NotI
Eco52I
HpaI
MunI

FIGURE 22

19 Apr 1999 Sequence Data Page 1

Molecule: p43msENC-AT, 6142 bps DNA Circular
Description: Ligation of inverted msEnhancer into p43-AAT*
File Name: p43smENC-AT.cm5, dated 19 Apr 1999
Printed: 1-6142 bps (Full), format Single Strand

```
   1 gggggggggg gggggggttg gccactccct ctctgcgcgc tcgctcgctc
  51 actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc
 101 ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca
 151 ctaggggttc ctagatctga cacccaaata tggcctgggg tgaggaatgg
 201 tgccgtcgcc atatttgggt gtccaccatt cctcaccgct ctaaaaataa
 251 ctcccgggag ttatttttag agcgccaaca cctgctgcct gcccaccatt
 301 cctcaccgct ctaaaaataa ctccccacca ttcctcaccc gtcgccatat
 351 ttgggtgtcg tgaggaatgg tgagatcttc aatattggcc attagccata
 401 ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca
 451 tacgttgtat ctatatcata atatgtacat ttatattggc tcatgtccaa
 501 tatgaccgcc atgttggcat tgattattga ctagttatta atagtaatca
 551 attacggggt cattagttca tagcccatat atggagttcc gcgttacata
 601 acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat
 651 tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc
 701 cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt
 751 acatcaagtg tatcatatgc caagtccgcc ccctattgac gtcaatgacg
 801 gtaaatggcc cgcctggcat tatgcccagt acatgacctt acgggacttt
 851 cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat
 901 gcggttttgg cagtacacca atgggcgtgg atagcggttt gactcacggg
 951 gatttccaag tctccacccc attgacgtca atgggagttt gttttggcac
1001 caaaatcaac gggactttcc aaaatgtcgt aataaccccg ccccgttgac
1051 gcaaatgggc ggtaggcgtg tacggtggga ggtctatata agcagagctc
1101 gtttagtgaa ccgtcagatc actagaagct ttattgcggt agtttatcac
1151 agttaaattg ctaacgcagt cagtgcttct gacacaacag tctcgaactt
1201 aagctgcaga agttggtcgt gaggcactgg gcaggtaagt atcaaggtta
1251 caagacaggt ttaaggagac caatagaaac tgggcttgtc gagacagaga
1301 agactcttgc gtttctgata ggcacctatt ggtcttactg acatccactt
1351 tgcctttctc tccacaggtg tccactccca gttcaattac agctcttaag
1401 gctagagtac ttaatacgac tcactatagg ctagaactag tggatccccc
1451 gggctgcagg aattcgatat caagcttggg gattttcagg caccaccact
1501 gacctgggac agtgaatcga caatgccgtc ttctgtctcg tggggcatcc
1551 tcctgctggc aggcctgtgc tgcctggtcc ctgtctccct ggctgaggat
1601 ccccagggag atgctgccca gaagacagat acatcccacc atgatcagga
1651 tcacccaacc ttcaacaaga tcacccccaa cctggctgag ttcgccttca
1701 gcctataccg ccagctggca caccagtcca acagcaccaa tatcttcttc
1751 tccccagtga gcatcgctac agcctttgca atgctctccc tggggaccaa
1801 ggctgacact cacgatgaaa tcctggaggg cctgaatttc aacctcacgg
1851 agattccgga ggctcagatc catgaaggct tccaggaact cctccgtacc
1901 ctcaaccagc cagacagcca gctccagctg accaccggca atggcctgtt
1951 cctcagcgag ggcctgaagc tagtggataa gtttttggag gatgttaaaa
2001 agttgtacca ctcagaagcc ttcactgtca acttcgggga caccgaagag
2051 gccaagaaac agatcaacga ttacgtggag aagggtactc aagggaaaat
2101 tgtggatttg gtcaaggagc ttgacagaga cacagttttt gctctggtga
2151 attacatctt ctttaaaggc aaatgggaga gacccttga agtcaaggac
2201 accgaggaag aggacttcca cgtggaccag gtgaccaccg tgaaggtgcc
2251 tatgatgaag cgtttaggca tgtttaacat ccagcactgt aagaagctgt
2301 ccagctgggt gctgctgatg aaatacctgg gcaatgccac cgccatcttc
2351 ttcctgcctg atgaggggaa actacagcac ctggaaaatg aactcaccca
2401 cgatatcatc accaagttcc tggaaaatga agacagaagg tctgccagct
```

FIGURE 22A

```
2451  tacatttacc caaactgtcc attactggaa cctatgatct gaagagcgtc
2501  ctgggtcaac tgggcatcac taaggtcttc agcaatgggg ctgacctctc
2551  cggggtcaca gaggaggcac ccctgaagct ctccaaggcc gtgcataagg
2601  ctgtgctgac catcgacgag aaagggactg aagctgctgg ggccatgttt
2651  ttagaggcca tacccatgtc tatccccccc gaggtcaagt tcaacaaacc
2701  ctttgtcttc ttaatgattg aacaaaatac caagtctccc ctcttcatgg
2751  gaaaagtggt gaatcccacc caaaaataac tgcctctcgc tcctcaaccc
2801  ctcccctcca tccctggccc cctccctgga tgacattaaa gaagggttga
2851  gctggtaacc ccccccccccc ctgcaggggc cctcgacccg ggcggccgct
2901  tcgagcagac atgataagat acattgatga gtttggacaa accacaacta
2951  gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct
3001  ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg
3051  cattcatttt atgtttcagg ttcaggggga gatgtgggag gtttttttaaa
3101  gcaagtaaaa cctctacaaa tgtggtaaaa tcgataagga tctaggaacc
3151  cctagtgatg gagttggcca ctccctctct gcgcgctcgc tcgctcactg
3201  aggccgcccg ggcaaagccc gggcgtcggg cgacctttgg tcgcccggcc
3251  tcagtgagcg agcgagcgcg cagagaggga gtggccaacc ccccccccccc
3301  ccccccctgca gcctggcgta atagcgaaga ggcccgcacc gatcgccctt
3351  cccaacagtt gcgtagcctg aatggcgaat ggcgcgacgc gccctgtagc
3401  ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac
3451  acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc
3501  tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctcccct
3551  ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga
3601  ttagggtgat ggttcacgta gtgggccatc gccctgatag acggtttttc
3651  gccctttgac gttggagtcc acgttcttta atagtggact cttgttccaa
3701  actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg
3751  gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa
3801  aatttaacgc gaattttaac aaaatattaa cgtttacaat ttcctgatgc
3851  ggtatttttct ccttacgcat ctgtgcggta tttcacaccg catatggtgc
3901  actctcagta caatctgctc tgatgccgca tagttaagcc agccccgaca
3951  cccgccaaca cccgctgacg cgccctgacg ggcttgtctg ctcccggcat
4001  ccgcttacag acaagctgtg accgtctccg ggagctgcat gtgtcagagg
4051  ttttcaccgt catcaccgaa acgcgcgaga cgaaagggcc tcgtgatacg
4101  cctatttttta taggttaatg tcatgataat aatggtttct tagacgtcag
4151  gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc
4201  taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat
4251  gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg
4301  tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac
4351  ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg
4401  agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt
4451  ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta
4501  tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg
4551  ccgcatacac tattctcaga atgacttggt tgagtactca ccagtcacag
4601  aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgctgcc
4651  ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg
4701  aggaccgaag gagctaaccg ctttttttgca caacatgggg gatcatgtaa
4751  ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac
4801  gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact
4851  attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact
4901  ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg
4951  gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg
5001  cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag
5051  ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag
5101  atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca
5151  agtttactca tatatacttt agattgattt aaaacttcat ttttaattta
```

FIGURE 22B

p43msENC-AT


```
5201   aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct
5251   taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa
5301   aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa
5351   caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta
5401   ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa
5451   tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg
5501   tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct
5551   gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt
5601   accggataag gcgcagcggt cgggctgaac gggggttcg tgcacacagc
5651   ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag
5701   cattgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc
5751   ggtaagcggc agggtcggaa caggagagcg cacgagggag cttccagggg
5801   gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt
5851   gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa
5901   cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg
5951   ctcacatgtt ctttcctgcg ttatcccctg attctgtgga taaccgtatt
6001   accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg
6051   cagcgagtca gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc
6101   ctctccccgc gcgttggccg attcattaat gcagggctgc ag
```

FIGURE 22C

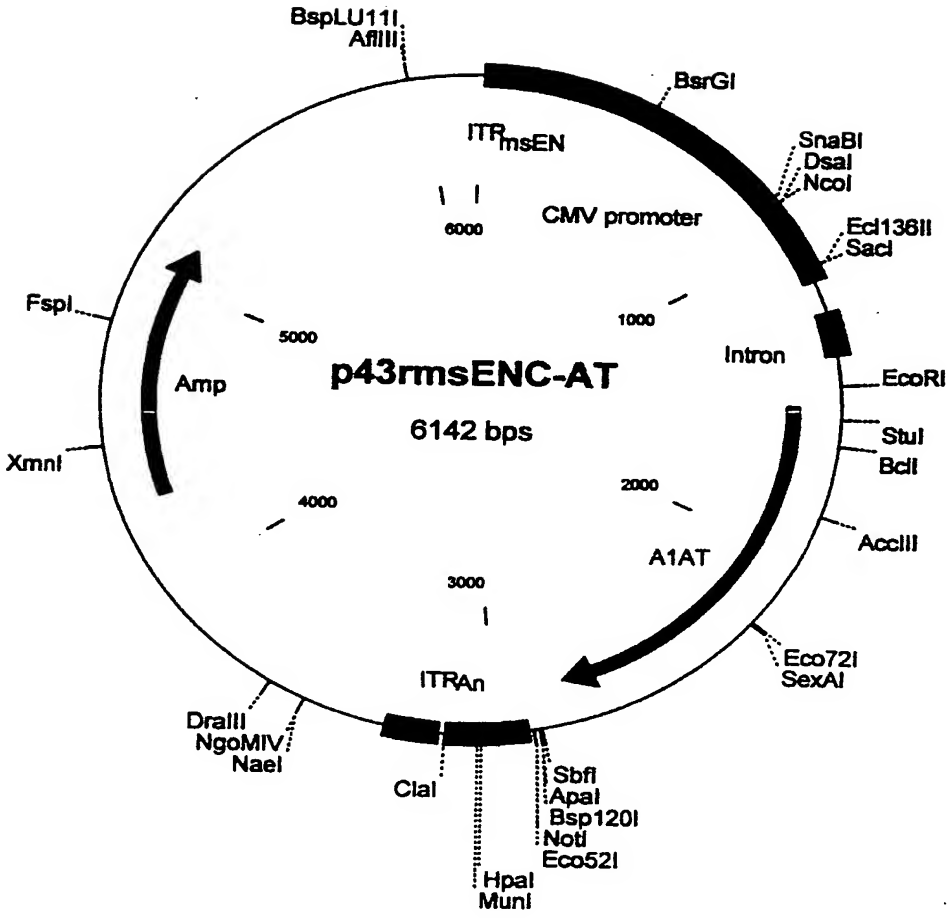

BspLU11I
AflIII
BsrGI
ITRmsEN
SnaBI
DsaI
NcoI
6000
CMV promoter
Ecl136II
SacI
1000
Intron
EcoRI
StuI
BclI
AccIII
2000
A1AT
Eco72I
SexAI
3000
ITRAn
SbfI
ApaI
Bsp120I
NotI
Eco52I
HpaI
MunI
ClaI
DraIII
NgoMIV
NaeI
4000
XmnI
5000
Amp
FspI
p43rmsENC-AT
6142 bps

FIGURE 23

19 Apr 1999

Sequence Data

Molecule: p43rmsENC-AT, 6142 bps DNA Circular
Description: Ligation of inverted msEnhancer into p43-AAT*
File Name: p43rmsENC-AT.cm5, dated 19 Apr 1999
Printed: 1-6142 bps (Full), format Single Strand

```
   1  gggggggggg gggggggttg gccactccct ctctgcgcgc tcgctcgctc
  51  actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc
 101  ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca
 151  ctaggggttc ctagatctga cacccaaata tggcctgggg tgaggaatgg
 201  tgccgtcgcc atatttgggt gtccaccatt cctcaccgct ctaaaaataa
 251  ctcccgggag ttatttttag agcgccaaca cctgctgcct gcccaccatt
 301  cctcaccgct ctaaaaataa ctccccacca ttcctcaccc gtcgccatat
 351  ttgggtgtcg tgaggaatgg tgagatcttc aatattggcc attagccata
 401  ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca
 451  tacgttgtat ctatatcata atatgtacat ttatattggc tcatgtccaa
 501  tatgaccgcc atgttggcat tgattattga ctagttatta atagtaatca
 551  attacggggt cattagttca tagcccatat atggagttcc gcgttacata
 601  acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat
 651  tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc
 701  cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt
 751  acatcaagtg tatcatatgc caagtccgcc ccctattgac gtcaatgacg
 801  gtaaatggcc cgcctggcat tatgcccagt acatgacctt acgggacttt
 851  cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat
 901  gcggttttgg cagtacacca atgggcgtgg atagcggttt gactcacggg
 951  gatttccaag tctccacccc attgacgtca atgggagttt gttttggcac
1001  caaaatcaac gggactttcc aaaatgtcgt aataaccccg ccccgttgac
1051  gcaaatgggc ggtaggcgtg tacggtggga ggtctatata agcagagctc
1101  gtttagtgaa ccgtcagatc actagaagct ttattgcggt agtttatcac
1151  agttaaattg ctaacgcagt cagtgcttct gacacaacag tctcgaactt
1201  aagctgcaga agttggtcgt gaggcactgg gcaggtaagt atcaaggtta
1251  caagacaggt ttaaggagac caatagaaac tgggcttgtc gagacagaga
1301  agactcttgc gtttctgata ggcacctatt ggtcttactg acatccactt
1351  tgcctttctc tccacaggtg tccactccca gttcaattac agctcttaag
1401  gctagagtac ttaatacgac tcactatagg ctagaactag tggatccccc
1451  gggctgcagg aattcgatat caagcttggg gattttcagg caccaccact
1501  gacctgggac agtgaatcga caatgccgtc ttctgtctcg tggggcatcc
1551  tcctgctggc aggcctgtgc tgcctggtcc ctgtctccct ggctgaggat
1601  ccccagggag atgctgccca gaagacagat acatcccacc atgatcagga
1651  tcacccaacc ttcaacaaga tcacccccaa cctggctgag ttcgccttca
1701  gcctataccg ccagctggca caccagtcca acagcaccaa tatcttcttc
1751  tccccagtga gcatcgctac agcctttgca atgctctccc tggggaccaa
1801  ggctgacact cacgatgaaa tcctggaggg cctgaatttc aacctcacgg
1851  agattccgga ggctcagatc catgaaggct tccaggaact cctccgtacc
1901  ctcaaccagc cagacagcca gctccagctg accaccggca atggcctgtt
1951  cctcagcgag ggcctgaagc tagtggataa gtttttggag gatgttaaaa
2001  agttgtacca ctcagaagcc ttcactgtca acttcgggga caccgaagag
2051  gccaagaaac agatcaacga ttacgtggag aagggtactc aagggaaaat
2101  tgtggatttg gtcaaggagc ttgacagaga cacagttttt gctctggtga
2151  attacatctt ctttaaaggc aaatgggaga gacccttga agtcaaggac
2201  accgaggaag aggacttcca cgtggaccag gtgaccaccg tgaaggtgcc
2251  tatgatgaag cgtttaggca tgtttaacat ccagcactgt aagaagctgt
2301  ccagctgggt gctgctgatg aaatacctgg gcaatgccac cgccatcttc
2351  ttcctgcctg atgaggggaa actacagcac ctggaaaatg aactcaccca
2401  cgatatcatc accaagttcc tggaaaatga agacagaagg tctgccagct
```

FIGURE 23A

p43rmsENC-AT


```
2451  tacatttacc caaactgtcc attactggaa cctatgatct gaagagcgtc
2501  ctgggtcaac tgggcatcac taaggtcttc agcaatgggg ctgacctctc
2551  cggggtcaca gaggaggcac ccctgaagct ctccaaggcc gtgcataagg
2601  ctgtgctgac catcgacgag aaagggactg aagctgctgg ggccatgttt
2651  ttagaggcca tacccatgtc tatccccccc gaggtcaagt tcaacaaacc
2701  ctttgtcttc ttaatgattg aacaaaatac caagtctccc ctcttcatgg
2751  gaaaagtggt gaatcccacc caaaaataac tgcctctcgc tcctcaaccc
2801  ctcccctcca tccctggccc cctccctgga tgacattaaa gaagggttga
2851  gctggtaacc ccccccccccc ctgcaggggc cctcgacccg ggcggccgct
2901  tcgagcagac atgataagat acattgatga gtttggacaa accacaacta
2951  gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct
3001  ttatttgtaa ccattataag ctgcaataaa caagttaaca acaacaattg
3051  cattcatttt atgtttcagg ttcaggggga gatgtgggag gtttttaaa
3101  gcaagtaaaa cctctacaaa tgtggtaaaa tcgataagga tctaggaacc
3151  cctagtgatg gagttggcca ctccctctct gcgcgctcgc tcgctcactg
3201  aggccgcccg ggcaaagccc gggcgtcggg cgacctttgg tcgcccggcc
3251  tcagtgagcg agcgagcgcg cagagaggga gtggccaacc ccccccccccc
3301  ccccctgca gcctggcgta atagcgaaga ggcccgcacc gatcgccctt
3351  cccaacagtt gcgtagcctg aatggcgaat ggcgcgacgc gccctgtagc
3401  ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac
3451  acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc
3501  tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg ggggctccct
3551  ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga
3601  ttagggtgat ggttcacgta gtgggccatc gccctgatag acggtttttc
3651  gccctttgac gttggagtcc acgttcttta atagtggact cttgttccaa
3701  actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg
3751  gattttgccg atttcggcct attggttaaa aaatgagctg atttaacaaa
3801  aatttaacgc gaattttaac aaaatattaa cgtttacaat ttcctgatgc
3851  ggtatttttct ccttacgcat ctgtgcggta tttcacaccg catatggtgc
3901  actctcagta caatctgctc tgatgccgca tagttaagcc agccccgaca
3951  cccgccaaca cccgctgacg cgccctgacg ggcttgtctg ctcccggcat
4001  ccgcttacag acaagctgtg accgtctccg ggagctgcat gtgtcagagg
4051  ttttcaccgt catcaccgaa acgcgcgaga cgaaagggcc tcgtgatacg
4101  cctatttttta taggttaatg tcatgataat aatggtttct tagacgtcag
4151  gtggcacttt tcggggaaat gtgcgcggaa cccctatttg tttatttttc
4201  taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat
4251  gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg
4301  tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac
4351  ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg
4401  agtgggttac atcgaactgg atctcaacag cggtaagatc cttgagagtt
4451  ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta
4501  tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg
4551  ccgcatacac tattctcaga atgacttggt tgagtactca ccagtcacag
4601  aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgctgcc
4651  ataaccatga gtgataacac tgcggccaac ttacttctga caacgatcgg
4701  aggaccgaag gagctaaccg ctttttttgca caacatgggg gatcatgtaa
4751  ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac
4801  gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact
4851  attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact
4901  ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg
4951  gctggctggt ttattgctga taaatctgga gccggtgagc gtgggtctcg
5001  cggtatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag
5051  ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag
5101  atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca
5151  agtttactca tatatacttt agattgattt aaaacttcat ttttaattta
```

FIGURE 23B

p43rmsENC-AT


```
5201   aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct
5251   taacgtgagt tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa
5301   aggatcttct tgagatcctt tttttctgcg cgtaatctgc tgcttgcaaa
5351   caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta
5401   ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa
5451   tactgtcctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg
5501   tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct
5551   gccagtggcg ataagtcgtg tcttaccggg ttggactcaa gacgatagtt
5601   accggataag gcgcagcggt cgggctgaac ggggggttcg tgcacacagc
5651   ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag
5701   cattgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc
5751   ggtaagcggc agggtcggaa caggagagcg cacgagggag cttccagggg
5801   gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt
5851   gagcgtcgat ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa
5901   cgccagcaac gcggcctttt tacggttcct ggccttttgc tggccttttg
5951   ctcacatgtt ctttcctgcg ttatcccctg attctgtgga taaccgtatt
6001   accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg
6051   cagcgagtca gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc
6101   ctctccccgc gcgttggccg attcattaat gcagggctgc ag
```

FIGURE 23C

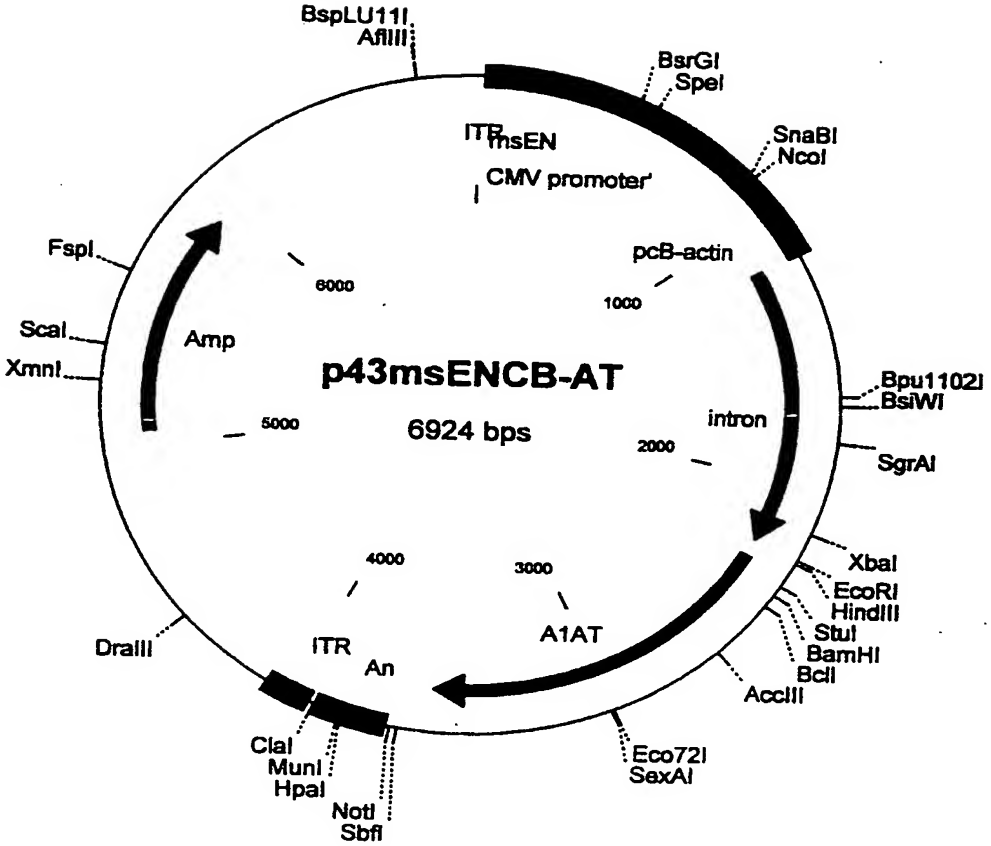

BspLU11I
AflIII
BsrGI
SpeI
SnaBI
NcoI
ITR
msEN
CMV promoter
pcB-actin
1000
intron
2000
SgrAI
Bpu1102I
BsiWI
XbaI
EcoRI
HindIII
StuI
BamHI
BclI
AccIII
Eco72I
SexAI
3000
A1AT
4000
ITR
An
ClaI
MunI
HpaI
NotI
SbfI
DraIII
FspI
ScaI
XmnI
Amp
6000
5000
p43msENCB-AT
6924 bps

FIGURE 24

19 Apr 1999

Sequence Data

Molecule: p43msENCB-AT, 6924 bps DNA Circular
Description: Ligation of msEnhacer into p43CB-AT*
File Name: p43msENCB-AT.cm5, dated 19 Apr 1999
Printed: 1-6924 bps (Full), format Single Strand

```
   1  ggggggggag ggggggttg gccactccct ctctgcgcgc tcgctcgctc
  51  actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc
 101  ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca
 151  ctaggggttc ctagatctca ccattcctca cgacacccaa atatggcgac
 201  gggtgaggaa tggtggggag ttatttttag agcggtgagg aatggtgggc
 251  aggcagcagg tgttggcgct ctaaaaataa ctcccgggag ttatttttag
 301  agcggtgagg aatggtggac acccaaatat ggcgacggca ccattcctca
 351  ccccaggcca tatttgggtg tcagatcttc aatattggcc attagccata
 401  ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca
 451  tacgttgtat ctatatcata atatgtacat ttatattggc tcatgtccaa
 501  tatgaccgcc atgttggcat tgattattga ctagttatta atagtaatca
 551  attacggggt cattagttca tagcccatat atggagttcc gcgttacata
 601  acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat
 651  tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc
 701  cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt
 751  acatcaagtg tatcatatgc caagtccgcc ccctattgac gtcaatgacg
 801  gtaaatggcc cgcctggcat tatgcccagt acatgacctt acgggacttt
 851  cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtcga
 901  ggtgagcccc acgttctgct tcactctccc catctccccc ccctccccac
 951  ccccaatttt gtatttattt atttttaat tattttgtgc agcgatgggg
1001  gcgggggggg gggggggcg cgcgccaggc ggggcggggc ggggcgaggg
1051  gcggggcggg gcgaggcgga gaggtgcggc ggcagccaat cagagcggcg
1101  cgctccgaaa gtttcctttt atggcgaggc ggcggcggcg gcggccctat
1151  aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc tgccttcgcc
1201  ccgtgccccg ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg
1251  accgcgttac tcccacaggt gagcgggcgg gacggccctt ctcctccggg
1301  ctgtaattag cgcttggttt aatgacggct tgtttctttt ctgtggctgc
1351  gtgaaagcct tgaggggctc cgggagggcc ctttgtgcgg gggggagcgg
1401  ctcggggggt gcgtgcgtgt gtgtgtgcgt ggggagcgcc gcgtgcggcc
1451  cgcgctgccc ggcggctgtg agcgctgcgg gcgcggcgcg gggctttgtg
1501  cgctccgcag tgtgcgcgag gggagcgcgg ccgggggcgg tgccccgcgg
1551  tgcgggggcg gctgcgaggg gaacaaaggc tgcgtgcggg gtgtgtgcgt
1601  gggggggtga gcaggggtg tgggcgcggc ggtcgggctg taacccccccc
1651  ctgcaccccc ctccccgagt tgctgagcac ggcccggctt cgggtgcggg
1701  gctccgtacg gggcgtggcg cggggctcgc cgtgccgggc gggggtggc
1751  ggcaggtggg ggtgccgggc ggggcggggc cgcctcgggc cggggagggc
1801  tcgggggagg ggcgcggcgg ccccggagc gccggcggct gtcgaggcgc
1851  ggcgagccgc agccattgcc ttttatggta atcgtgcgag agggcgcagg
1901  gacttccttt gtcccaaatc tgtgcggagc cgaaatctgg gaggcgccgc
1951  cgcacccccc ctagcgggcg cggggcgaag cggtgcggcg ccggcaggaa
2001  ggaaatgggc ggggagggcc ttcgtgcgtc gccgcgccgc cgtcccccttc
2051  tccctctcca gcctcggggc tgtccgcggg gggacggctg ccttcggggg
2101  ggacggggca gggcggggtt cggcttctgg cgtgtgaccg gcggctctag
2151  agcctctgct aaccatgttc atgccttctt ctttttccta cagctcctgg
2201  gcaacgtgct ggttattgtg ctgtctcatc attttggcaa agaattcgat
2251  atcaagcttg gggattttca ggcaccacca ctgacctggg acagtgaatc
2301  gacaatgccg tcttctgtct cgtggggcat cctcctgctg gcaggcctgt
2351  gctgcctggt ccctgtctcc ctggctgagg atccccaggg agatgctgcc
2401  cagaagacag atacatccca ccatgatcag gatcacccaa ccttcaacaa
```

FIGURE 24A

p43msENCB-AT


```
2451  gatcaccccc aacctggctg agttcgcctt cagcctatac cgccagctgg
2501  cacaccagtc caacagcacc aatatcttct tctccccagt gagcatcgct
2551  acagcctttg caatgctctc cctggggacc aaggctgaca ctcacgatga
2601  aatcctggag ggcctgaatt tcaacctcac ggagattccg gaggctcaga
2651  tccatgaagg cttccaggaa ctcctccgta ccctcaacca gccagacagc
2701  cagctccagc tgaccaccgg caatggcctg ttcctcagcg agggcctgaa
2751  gctagtggat aagtttttgg aggatgttaa aaagttgtac cactcagaag
2801  ccttcactgt caacttcggg gacaccgaag aggccaagaa acagatcaac
2851  gattacgtgg agaagggtac tcaagggaaa attgtggatt tggtcaagga
2901  gcttgacaga gacacagttt ttgctctggt gaattacatc ttctttaaag
2951  gcaaatggga gagacccttt gaagtcaagg acaccgagga agaggacttc
3001  cacgtggacc aggtgaccac cgtgaaggtg cctatgatga agcgtttagg
3051  catgtttaac atccagcact gtaagaagct gtccagctgg gtgctgctga
3101  tgaaatacct gggcaatgcc accgccatct tcttcctgcc tgatgagggg
3151  aaactacagc acctggaaaa tgaactcacc cacgatatca tcaccaagtt
3201  cctggaaaat gaagacagaa ggtctgccag cttacattta cccaaactgt
3251  ccattactgg aacctatgat ctgaagagcg tcctgggtca actgggcatc
3301  actaaggtct tcagcaatgg ggctgacctc tccggggtca cagaggaggc
3351  acccctgaag ctctccaagg ccgtgcataa ggctgtgctg accatcgacg
3401  agaaagggac tgaagctgct ggggccatgt ttttagaggc catacccatg
3451  tctatccccc ccgaggtcaa gttcaacaaa ccctttgtct tcttaatgat
3501  tgaacaaaat accaagtctc ccctcttcat gggaaaagtg gtgaatccca
3551  cccaaaaata actgcctctc gctcctcaac ccctcccctc catccctggc
3601  cccctccctg gatgacatta aagaagggtt gagctggtaa ccccccccccc
3651  ccctgcaggg gccctcgacc cgggcggccg cttcgagcag acatgataag
3701  atacattgat gagtttggac aaaccacaac tagaatgcag tgaaaaaaat
3751  gctttatttg tgaaatttgt gatgctattg ctttatttgt aaccattata
3801  agctgcaata aacaagttaa caacaacaat tgcattcatt ttatgtttca
3851  ggttcagggg gagatgtggg aggtttttta aagcaagtaa aacctctaca
3901  aatgtggtaa aatcgataag gatctaggaa cccctagtga tggagttggc
3951  cactccctct ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc
4001  ccgggcgtcg ggcgaccttt ggtcgcccgg cctcagtgag cgagcgagcg
4051  cgcagagagg gagtggccaa ccccccccccc ccccccccctg cagcctggcg
4101  taatagcgaa gaggcccgca ccgatcgccc ttcccaacag ttgcgtagcc
4151  tgaatggcga atggcgcgac gcgccctgta gcggcgcatt aagcgcggcg
4201  ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc
4251  gcccgctcct ttcgctttct tcccttcctt tctcgccacg ttcgccggct
4301  ttccccgtca agctctaaat cgggggctcc ctttagggtt ccgatttagt
4351  gctttacggc acctcgaccc caaaaaactt gattagggtg atggttcacg
4401  tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt
4451  ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac
4501  cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc
4551  ctattggtta aaaaatgagc tgatttaaca aaaatttaac gcgaatttta
4601  acaaaatatt aacgtttaca atttcctgat gcggtatttt ctccttacgc
4651  atctgtgcgg tatttcacac cgcatatggt gcactctcag tacaatctgc
4701  tctgatgccg catagttaag ccagccccga cacccgccaa cacccgctga
4751  cgcgccctga cgggcttgtc tgctcccggc atccgcttac agacaagctg
4801  tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg
4851  aaacgcgcga gacgaaaggg cctcgtgata cgcctatttt tataggttaa
4901  tgtcatgata ataatggttt cttagacgtc aggtggcact tttcggggaa
4951  atgtgcgcgg aacccctatt tgtttatttt tctaaataca ttcaaatatg
5001  tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa
5051  aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt
5101  ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa
5151  gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact
```

FIGURE 24B

p43msENCB-AT


```
5201  ggatctcaac agcggtaaga tccttgagag ttttcgcccc gaagaacgtt
5251  ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc
5301  cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca
5351  gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg
5401  gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac
5451  actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac
5501  cgcttttttg cacaacatgg gggatcatgt aactcgcctt gatcgttggg
5551  aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg
5601  cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact
5651  tactctagct tcccggcaac aattaataga ctggatggag gcggataaag
5701  ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct
5751  gataaatctg gagccggtga gcgtgggtct cgcggtatca ttgcagcact
5801  ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga
5851  gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc
5901  tcactgatta agcattggta actgtcagac caagtttact catatatact
5951  ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga
6001  tccttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc
6051  cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc
6101  ttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac
6151  cagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag
6201  gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta
6251  gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc
6301  tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg
6351  tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg
6401  gtcgggctga acgggggtt cgtgcacaca gcccagcttg gagcgaacga
6451  cctacaccga actgagatac ctacagcgtg agcattgaga aagcgccacg
6501  cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg
6551  aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt
6601  atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga
6651  tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt
6701  tttacggttc ctggcctttt gctggccttt tgctcacatg ttctttcctg
6751  cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgagct
6801  gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga
6851  ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc
6901  cgattcatta atgcagggct gcag
```

FIGURE 24C

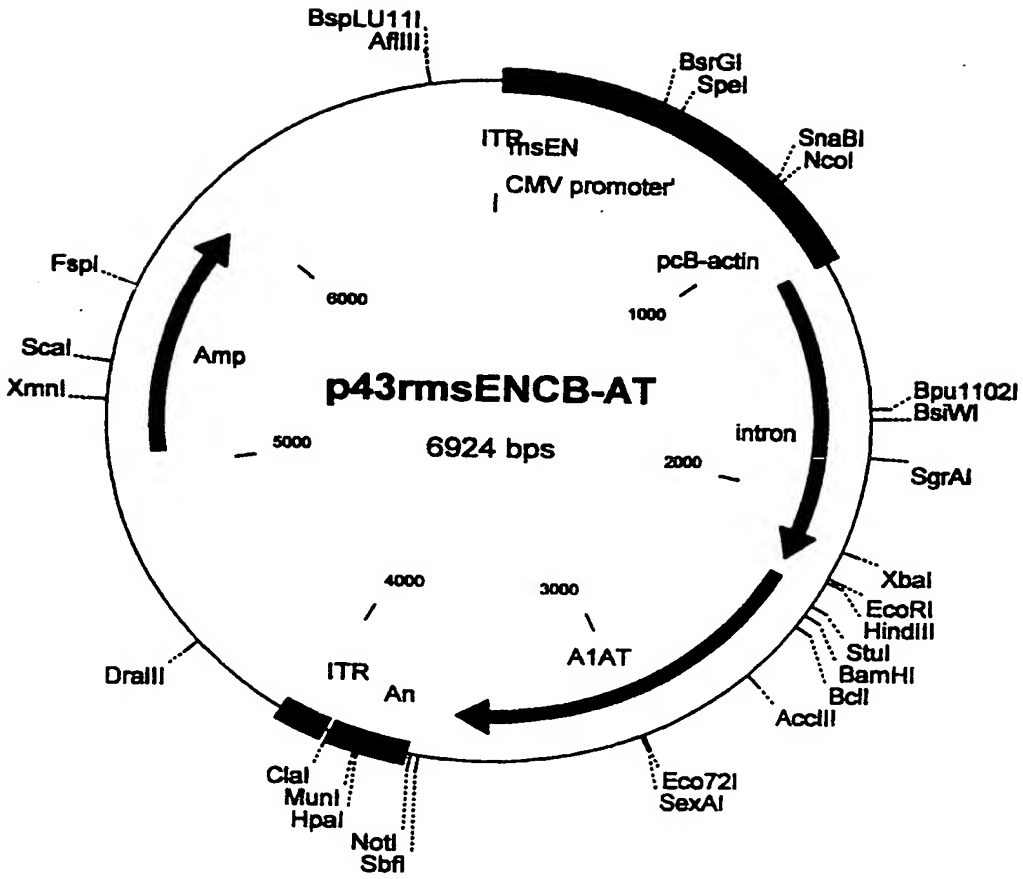

BspLU11I
AflIII
BsrGI
SpeI
SnaBI
NcoI
ITR
rmsEN
CMV promoter
pcB-actin
1000
2000
intron
SgrAI
Bpu1102I
BsiWI
XbaI
EcoRI
HindIII
StuI
BamHI
BclI
AccIII
Eco72I
SexAI
3000
A1AT
4000
ITR
An
ClaI
MunI
HpaI
NotI
SbfI
DraIII
5000
Amp
6000
FspI
ScaI
XmnI
p43rmsENCB-AT
6924 bps

FIGURE 25

Molecule: p43rmsENCB-AT, 6924 bps DNA Circular
Description: Ligation of inverted msEnhacer into p43CB-AT*
File Name: p43rmsCB-AT.cm5, dated 19 Apr 1999
Printed: 1-6924 bps (Full), format Single Strand

```
   1 gggggggggg gggggggttg gccactccct ctctgcgcgc tcgctcgctc
  51 actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc
 101 ggcctcagtg agcgagcgag cgcgcagaga gggagtggcc aactccatca
 151 ctaggggttc ctagatctga cacccaaata tggcctgggg tgaggaatgg
 201 tgccgtcgcc atatttgggt gtccaccatt cctcaccgct ctaaaaataa
 251 ctcccgggag ttatttttag agcgccaaca cctgctgcct gcccaccatt
 301 cctcaccgct ctaaaaataa ctccccacca ttcctcaccc gtcgccatat
 351 ttgggtgtcg tgaggaatgg tgagatcttc aatattggcc attagccata
 401 ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca
 451 tacgttgtat ctatatcata atatgtacat ttatattggc tcatgtccaa
 501 tatgaccgcc atgttggcat tgattattga ctagttatta atagtaatca
 551 attacggggt cattagttca tagcccatat atggagttcc gcgttacata
 601 acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat
 651 tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc
 701 cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt
 751 acatcaagtg tatcatatgc caagtccgcc ccctattgac gtcaatgacg
 801 gtaaatggcc cgcctggcat tatgcccagt acatgacctt acgggacttt
 851 cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtcga
 901 ggtgagcccc acgttctgct tcactctccc catctccccc ccctccccac
 951 ccccaatttt gtatttattt atttttaat tattttgtgc agcgatgggg
1001 gcgggggggg gggggggcgg cgcgccaggc ggggcggggc ggggcgaggg
1051 gcggggcggg gcgaggcgga gaggtgcggc ggcagccaat cagagcggcg
1101 cgctccgaaa gtttcctttt atggcgaggc ggcggcggcg gcggccctat
1151 aaaaagcgaa gcgcgcggcg ggcgggagtc gctgcgacgc tgccttcgcc
1201 ccgtgccccg ctccgccgcc gcctcgcgcc gcccgccccg gctctgactg
1251 accgcgttac tcccacaggt gagcgggcgg gacggcccct ctcctccggg
1301 ctgtaattag cgcttggttt aatgacggct tgtttctttt ctgtggctgc
1351 gtgaaagcct tgaggggctc cgggagggcc ctttgtgcgg gggggagcgg
1401 ctcgggggggt gcgtgcgtgt gtgtgtgcgt ggggagcgcc gcgtgcggcc
1451 cgcgctgccc ggcggctgtg agcgctgcgg gcgcggcgcg gggctttgtg
1501 cgctccgcag tgtgcgcgag gggagcgcgg ccgggggcgg tgccccgcgg
1551 tgcgggggggg gctgcgaggg gaacaaaggc tgcgtgcggg gtgtgtgcgt
1601 gggggggtga gcaggggggtg tgggcgcggc ggtcgggctg taacccccccc
1651 ctgcaccccc ctccccgagt tgctgagcac ggcccggctt cgggtgcggg
1701 gctccgtacg gggcgtggcg cggggctcgc cgtgccgggc gggggggtggc
1751 ggcaggtggg ggtgccgggc ggggcggggc cgcctcgggc cggggagggc
1801 tcgggggagg ggcgcggcgg cccccggagc gccggcggct gtcgaggcgc
1851 ggcgagccgc agccattgcc ttttatggta atcgtgcgag agggcgcagg
1901 gacttccttt gtcccaaatc tgtgcggagc cgaaatctgg gaggcgccgc
1951 cgcacccccct ctagcgggcg cggggcgaag cggtgcggcg ccggcaggaa
2001 ggaaatgggc ggggagggcc ttcgtgcgtc gccgcgccgc cgtcccccttc
2051 tccctctcca gcctcggggc tgtccgcggg gggacggctg ccttcggggg
2101 ggacggggca gggcggggtt cggcttctgg cgtgtgaccg gcggctctag
2151 agcctctgct aaccatgttc atgccttctt ctttttccta cagctcctgg
2201 gcaacgtgct ggttattgtg ctgtctcatc attttggcaa agaattcgat
2251 atcaagcttg gggattttca ggcaccacca ctgacctggg acagtgaatc
2301 gacaatgccg tcttctgtct cgtggggcat cctcctgctg gcaggcctgt
2351 gctgcctggt ccctgtctcc ctggctgagg atccccaggg agatgctgcc
2401 cagaagacag atacatccca ccatgatcag gatcacccaa ccttcaacaa
```

FIGURE 25A

```
2451  gatcaccccc aacctggctg agttcgcctt cagcctatac cgccagctgg
2501  cacaccagtc caacagcacc aatatcttct tctccccagt gagcatcgct
2551  acagcctttg caatgctctc cctggggacc aaggctgaca ctcacgatga
2601  aatcctggag ggcctgaatt tcaacctcac ggagattccg gaggctcaga
2651  tccatgaagg cttccaggaa ctcctccgta ccctcaacca gccagacagc
2701  cagctccagc tgaccaccgg caatggcctg ttcctcagcg agggcctgaa
2751  gctagtggat aagtttttgg aggatgttaa aaagttgtac cactcagaag
2801  ccttcactgt caacttcggg gacaccgaag aggccaagaa acagatcaac
2851  gattacgtgg agaagggtac tcaagggaaa attgtggatt tggtcaagga
2901  gcttgacaga gacacagttt ttgctctggt gaattacatc ttctttaaag
2951  gcaaatggga gagacccttt gaagtcaagg acaccgagga agaggacttc
3001  cacgtggacc aggtgaccac cgtgaaggtg cctatgatga agcgtttagg
3051  catgtttaac atccagcact gtaagaagct gtccagctgg gtgctgctga
3101  tgaaatacct gggcaatgcc accgccatct tcttcctgcc tgatgagggg
3151  aaactacagc acctggaaaa tgaactcacc cacgatatca tcaccaagtt
3201  cctggaaaat gaagacagaa ggtctgccag cttacattta cccaaactgt
3251  ccattactgg aacctatgat ctgaagagcg tcctgggtca actgggcatc
3301  actaaggtct tcagcaatgg ggctgacctc tccggggtca cagaggaggc
3351  acccctgaag ctctccaagg ccgtgcataa ggctgtgctg accatcgacg
3401  agaaagggac tgaagctgct ggggccatgt ttttagaggc catacccatg
3451  tctatccccc ccgaggtcaa gttcaacaaa ccctttgtct tcttaatgat
3501  tgaacaaaat accaagtctc ccctcttcat gggaaaagtg gtgaatccca
3551  cccaaaaata actgcctctc gctcctcaac ccctcccctc catcctggc
3601  cccctccctg gatgacatta aagaagggtt gagctggtaa ccccccccccc
3651  ccctgcaggg gccctcgacc cgggcggccg cttcgagcag acatgataag
3701  atacattgat gagtttggac aaaccacaac tagaatgcag tgaaaaaaat
3751  gctttatttg tgaaatttgt gatgctattg ctttatttgt aaccattata
3801  agctgcaata aacaagttaa caacaacaat tgcattcatt ttatgtttca
3851  ggttcagggg gagatgtggg aggtttttta aagcaagtaa aacctctaca
3901  aatgtggtaa aatcgataag gatctaggaa cccctagtga tggagttggc
3951  cactccctct ctgcgcgctc gctcgctcac tgaggccgcc cgggcaaagc
4001  ccgggcgtcg ggcgaccttt ggtcgcccgg cctcagtgag cgagcgagcg
4051  cgcagagagg gagtggccaa ccccccccccc ccccccctg cagcctggcg
4101  taatagcgaa gaggcccgca ccgatcgccc ttcccaacag ttgcgtagcc
4151  tgaatggcga atggcgcgac gcgccctgta gcggcgcatt aagcgcggcg
4201  ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc
4251  gcccgctcct ttcgctttct tcccttcctt tctcgccacg ttcgccggct
4301  ttccccgtca agctctaaat cgggggctcc ctttagggtt ccgatttagt
4351  gctttacggc acctcgaccc caaaaaactt gattagggtg atggttcacg
4401  tagtgggcca tcgccctgat agacggtttt tcgccctttg acgttggagt
4451  ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac
4501  cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc
4551  ctattggtta aaaaatgagc tgatttaaca aaaatttaac gcgaatttta
4601  acaaaatatt aacgtttaca atttcctgat gcggtatttt ctccttacgc
4651  atctgtgcgg tatttcacac cgcatatggt gcactctcag tacaatctgc
4701  tctgatgccg catagttaag ccagccccga cacccgccaa cacccgctga
4751  cgcgccctga cgggcttgtc tgctcccggc atccgcttac agacaagctg
4801  tgaccgtctc cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg
4851  aaacgcgcga gacgaaaggg cctcgtgata cgcctatttt tataggttaa
4901  tgtcatgata ataatggttt cttagacgtc aggtggcact tttcggggaa
4951  atgtgcgcgg aacccctatt tgtttatttt tctaaataca ttcaaatatg
5001  tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa
5051  aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt
5101  ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa
5151  gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact
```

FIGURE 25B

```
5201  ggatctcaac agcggtaaga tccttgagag ttttcgcccc gaagaacgtt
5251  ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc
5301  cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca
5351  gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg
5401  gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac
5451  actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac
5501  cgcttttttg cacaacatgg gggatcatgt aactcgcctt gatcgttggg
5551  aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg
5601  cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact
5651  tactctagct tcccggcaac aattaataga ctggatggag gcggataaag
5701  ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct
5751  gataaatctg gagccggtga gcgtgggtct cgcggtatca ttgcagcact
5801  ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga
5851  gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc
5901  tcactgatta agcattggta actgtcagac caagtttact catatatact
5951  ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga
6001  tccttttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc
6051  cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc
6101  ttttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac
6151  cagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag
6201  gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta
6251  gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc
6301  tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg
6351  tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg
6401  gtcgggctga acgggggttt cgtgcacaca gcccagcttg gagcgaacga
6451  cctacaccga actgagatac ctacagcgtg agcattgaga aagcgccacg
6501  cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg
6551  aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt
6601  atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga
6651  tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt
6701  tttacggttc ctggcctttt gctggccttt tgctcacatg ttctttcctg
6751  cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgagct
6801  gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga
6851  ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc
6901  cgattcatta atgcagggct gcag
```

FIGURE 25C